Input–Output Behavior of ErbB Signaling Pathways as Revealed by a Mass Action Model Trained against Dynamic Data

Molecular Systems Biology 5; Article number 239; doi:10.1038/msb.2008.74
Citation: ***Molecular Systems Biology*** **5:239**

www.molecularsystemsbiology.com

molecular systems biology

Input–output behavior of ErbB signaling pathways as revealed by a mass action model trained against dynamic data

William W Chen[1,4], Birgit Schoeberl[2,4], Paul J Jasper[1,4], Mario Niepel[1], Ulrik B Nielsen[2], Douglas A Lauffenburger[3] and Peter K Sorger[1,*]

[1] Department of Systems Biology, Center for Cell Decision Processes, Harvard Medical School, Boston, MA, USA, [2] Merrimack Pharmaceuticals, Cambridge, MA, USA and [3] Department of Biological Engineering, Center for Cell Decision Processes, Massachusetts Institute of Technology, Cambridge, MA, USA
[4] These authors contributed equally to this work
* Corresponding author. Department of Systems Biology, Center for Cell Decision Processes, Harvard Medical School, Warren Alpert 438, 200 Longwood Avenue, Boston, MA 02115, USA. Tel.: +1 617 432 6901; Fax: +1 617 432 5012; E-mail: peter_sorger@hms.harvard.edu

Received 25.1.08; accepted 3.12.08

The ErbB signaling pathways, which regulate diverse physiological responses such as cell survival, proliferation and motility, have been subjected to extensive molecular analysis. Nonetheless, it remains poorly understood how different ligands induce different responses and how this is affected by oncogenic mutations. To quantify signal flow through ErbB-activated pathways we have constructed, trained and analyzed a mass action model of immediate-early signaling involving ErbB1–4 receptors (EGFR, HER2/Neu2, ErbB3 and ErbB4), and the MAPK and PI3K/Akt cascades. We find that parameter sensitivity is strongly dependent on the feature (e.g. ERK or Akt activation) or condition (e.g. EGF or heregulin stimulation) under examination and that this context dependence is informative with respect to mechanisms of signal propagation. Modeling predicts log-linear amplification so that significant ERK and Akt activation is observed at ligand concentrations far below the K_d for receptor binding. However, MAPK and Akt modules isolated from the ErbB model continue to exhibit switch-like responses. Thus, key system-wide features of ErbB signaling arise from nonlinear interaction among signaling elements, the properties of which appear quite different in context and in isolation.
Molecular Systems Biology 20 January 2009; doi:10.1038/msb.2008.74
Subject Categories: simulation and data analysis; signal transduction
Keywords: EGFR; ErbB; ODE model; parameter optimization; signal transduction

Introduction

The ErbB1–4 receptor tyrosine kinases (RTKs) and the signaling pathways they activate, control cell division, motility and survival and are among the best studied of all signal-transduction networks (Citri and Yarden, 2006). Hyperactive and aberrant ErbB signaling has been implicated in a wide variety of human cancers and is a frequent target of pharmaceutical intervention (Slamon, 1987; Rusch, 1993; Engelman, 2007). ErbB1–4 receptors bind as homo- and heterodimers to a family of 13 soluble and membrane-bound ligands and activate multiple downstream signaling pathways, including the mitogenic Ras-MAPK cascade and the pro-survival PI3K/Akt pathway. Complexity arises in ErbB networks from hetero- and homo-oligomerization among receptors having distinct ligand-binding properties and different intracellular binding partners, and from the multiplicity of intracellular proteins that interact with these receptors. A wide variety of anti-ErbB drugs are in use or development, including small-molecule ATP competitors (e.g. gefitinib and erlotinib) and therapeutic antibodies (e.g. cetuximab and trastuzumab), but most drugs exhibit significant patient-to-patient variation in efficacy (Slamon, 2001; Paez, 2004). The causes of this variation are under active investigation (Ono, 2004; Carey, 2006; Mulloy, 2007) and appear to be multi-factorial, even at the level of immediate-early signaling (Jasper, in preparation). Quantitative, network-level analysis is therefore needed to predict the consequences for signal transduction of changes in the activities and levels of multiple proteins. Toward this end, we have developed a computational model of the mammalian ErbB network that includes signaling from all four receptors and the ERK and Akt signal-transduction cascades.

ErbB receptors are single-pass type I transmembrane receptors with extracellular ligand-binding domains, intracellular

tyrosine kinase domains and cytoplasmic tails that act as signaling scaffolds. ErbB1 and ErbB4 are fully functional as ligand-dependent tyrosine kinases, but ErbB2 does not bind any known ligand, functioning instead as a dimerization-ready signal amplifier (Klapper, 1999). ErbB3 has a crippled kinase domain lacking catalytic activity (Guy, 1994) and therefore transduces signals only when phosphorylated by other ErbB receptors. The 13 known ErbB ligands can be divided into three groups: (i) those that bind specifically to ErbB1, such as epidermal growth factor (EGF), transforming growth factor alpha and amphiregulin, (ii) those that bind to both ErbB1 and ErbB4, including betacellulin, heparin-binding EGF, epigen and epiregulin and (iii) the neuregulins (NRGs), which fall into a subgroup comprising NRG1 (also known as GGF2, SMDF or HRG) and NRG2 that binds to heterodimers containing ErbB3 or ErbB4 and a second subgroup comprising NRG3/NRG4 that binds ErbB4 homodimers (Figure 1; Supplementary Figure 1A). Concomitant with ligand binding, ErbB receptors dimerize and undergo trans-phosphorylation on residues in their cytoplasmic tails thereby creating docking sites for SH2-containing adapter molecules, such as Shc, Grb2, Sos and PI3K. ErbB1 has at least 20 sites of tyrosine phosphorylation on its cytoplasmic tail, 12 of which are thought to partner with SH2-containing adapter proteins and enzymes (Schulze *et al*, 2005). Other ErbB receptors undergo equally extensive post-translational modification and adapter binding. Receptor-associated adapters activate Ras, ERK and Akt. Akt can also be activated in a Ras-independent manner through the direct binding of PI3K–p85 to multiple sites on ErbB3. Key effectors of ERK and Akt include the Elk-1, AP-1 and NF-κB transcription factors. In addition, ErbB receptors also activate protein kinase C, actin-based morphogenic responses and Jak–STAT signaling, but we have not yet modeled these pathways. Finally, the ErbB response is silenced by internalization and degradation of receptors, dephosphorylation by phosphatases and negative feedback from active ERK and Akt.

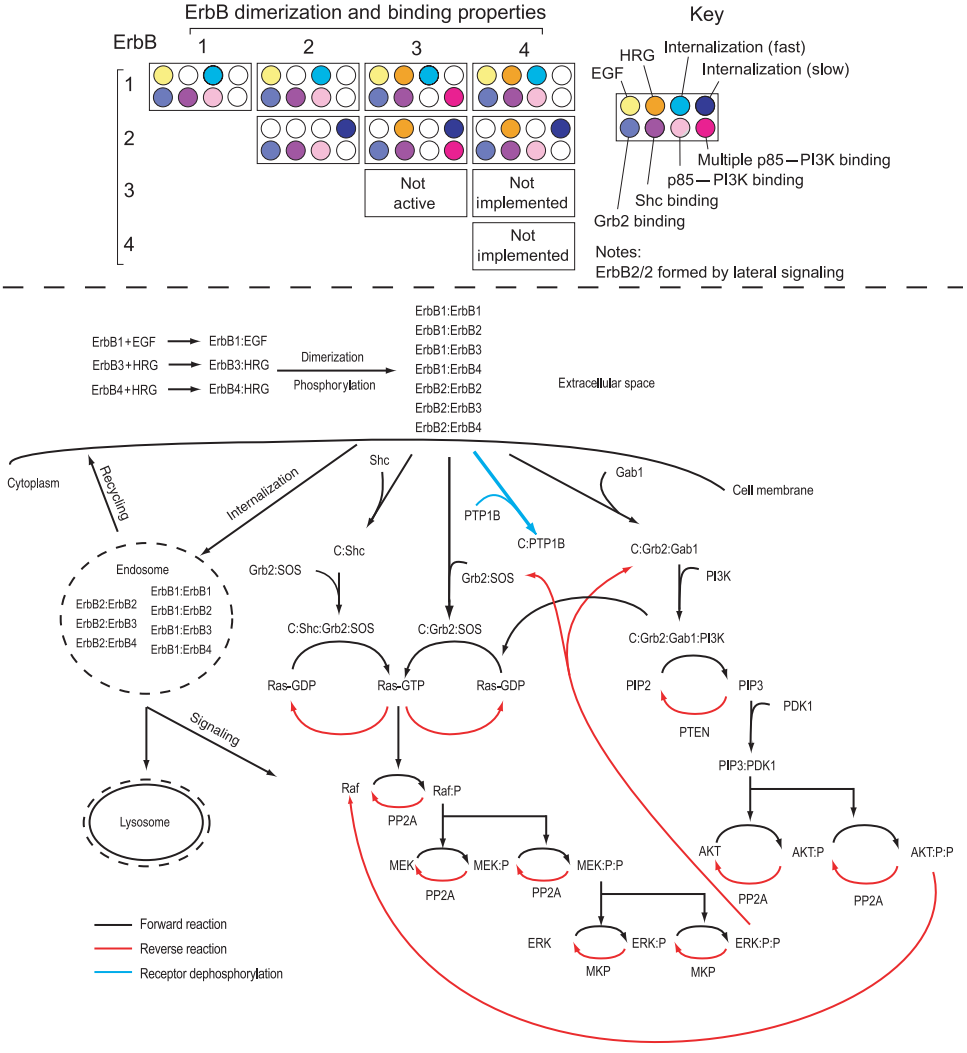


Figure 1 Simplified schematic representation of the ErbB model. Receptor interaction, internalization, recycling and activation of MAPK and PI3K/Akt pathways are shown. Phosphatase, Ras-GTP hydrolysis and feedback regulation are indicated by red arrows. Matrix shows properties of each receptor dimer relevant to the topology of the model; see Figure 1 in Supplementary information for further information. The prefix C on certain proteins denotes an abbreviated representation of multiple species in the model, e.g. C:Shc is composed of all receptor-bound Shc molecules, which may include ErbB1/1-Shc, ErbB1/2-Shc and so on. However, each of these Shc complexes is represented explicitly by one or more dynamic variables in the model.

The 'hour-glass' shape of RTK signaling networks (Citri and Yarden, 2006), in which a broad array of ligands and receptors funnel into a limited number of core signaling cascades (MAPK, PI3K–Akt, Jak–STAT and so on) to control a diverse set of transcription factors, raises fundamental questions as to how signals are transduced from the cell surface to the nucleus without loss of information. This question is particularly acute in the case of ErbB networks, in which multiple ligands control a diversity of physiological responses in a cell-specific manner, although nonetheless engaging a common set of signaling proteins. It seems probable that the ErbB network encodes information through changes in the stoichiometry and half-lives of post-translational modifications and the localization of interacting kinases, phosphatases and so on. To begin to understand these phenomena, we require a quantitative framework that encompasses the diversity of receptor hetero- and homo-oligomerization and at least some of the complexity of upstream ligands and downstream signaling. We also require an effective means to incorporate experimental data into the model in a manner that reproduces ligand- and cell-specific responses while taking into account uncertainty in the precise biochemistry of ErbB signaling and the inability of experiments to measure all relevant processes. Here, we present, as a substantial extension of our earlier work (Schoeberl *et al*, 2002), an ordinary differential equation (ODE) model of ERK and Akt regulation by two ErbB ligands and four ErbB receptors during the immediate-early phase of ligand-stimulated cell signaling. Relative to previously published models of ErbB signaling (Kholodenko *et al*, 1999; Hatakeyama, 2003; Resat *et al*, 2003; Hendriks, 2005; Sasagawa *et al*, 2005; Birtwistle *et al*, 2007), ours is more extensive in including all four receptors and two ligand classes, although nonetheless retaining the rigor of a mass action formulation based on elementary reactions. A recently published ErbB model by Birtwistle *et al* (2007) is also based on a kinetic representation of immediate-early signaling (from 0 to 30 min rather than 0 to 120 min as in our model), but aggregates species to reduce complexity. In contrast, we rely on elementary reactions throughout, albeit at the cost of more species and parameters. We consider parametric uncertainty and model non-identifiability explicitly and account for the fact that parameter sensitivity or robustness can only be interpreted in light of this uncertainty. Despite its non-identifiability, our model predicts experimentally verifiable system-wide features, such as variable amplification in receptor-activated enzymes as the basis of a very broad range in dose responsiveness.

Results

To construct a computational model of ErbB-mediated signaling, we extended our previous model (Schoeberl *et al*, 2002), which contained only the ErbB1 receptor, using information from models of other mammalian signal-transduction cascades and a large body of newly published data (Yarden and Sliwkowski, 2001; Citri and Yarden, 2006). Our primary goal was expanding from a single receptor to all four members of the ErbB family, thereby enabling analysis of receptor–receptor interaction and determinants of differential signaling by ligands such as EGF and HRG. The Immediate-early ErbB Reaction Model (IERMv1.0) in this paper was implemented in a deterministic, continuum approximation as a network of ODEs with temporal coverage from 0 to 120 min after ligand addition.

Seven ErbB hetero and homodimers that have been described in the literature were included in our model: ErbB1/1, ErbB1/2, ErbB1/3, ErbB1/4, ErbB2/2, ErbB2/3 and ErbB2/4. The majority of these dimers are activated by ligand binding, but several arise through a process of 'lateral signaling' (or secondary dimerization) in which dimers phosphorylated in a ligand-dependent manner dissociate into monomers that then reassociate with either phosphorylated or unphosphorylated monomers to create new homo- or hetero-dimers (Graus-Porta *et al*, 1997); in this way, active ErbB2/2 can form even though it is does not bind ligand. Because it has no kinase activity and is presumably inactive (Guy, 1994), ErbB3/3 was omitted from the model. ErbB3/4 and ErbB4/4 were also omitted because the cell lines we studied express ErbB3 and ErbB4 at substantially lower concentrations than ErbB1 and ErbB2 (receptor numbers for A431, H1666 and H3255 lines are reported in Supplementary Figure 2B). Each receptor dimer spawns a cascade of downstream reactions and omitting ErbB3/3, ErbB3/4 and ErbB4/4 reduces the total number of species in the model by $\sim$250, making numerical integration and parameter estimation significantly easier.

Model scope

Two critical decisions that arise during model design involve scope and level of detail. One point of view holds that models should include all known components and relevant phenomena so as to capture cellular biochemistry in as complete and realistic a manner as possible. However, as the number of proteins increases, so does the number of free parameters that govern their interaction, often in a nonlinear manner. A second point of view holds that models should be as simple as possible as long as they reproduce empirical findings; ideally, models should have fewer degrees of freedom than training data (Aldridge *et al*, 2006). However, strict application of this reasoning generates models in which biochemical processes are lumped together and the functions of individual proteins are difficult to discern. In common with previous studies (Kholodenko *et al*, 1999; Hatakeyama, 2003; Hendriks, 2005), we therefore chose to model ErbB signaling at an intermediate resolution in which key proteins were represented explicitly, but without accounting for all assembly intermediates. General cellular processes, such as protein degradation and endocytosis, were represented without molecular detail as simple first- or second-order reactions, the rates of which were estimated by optimization (see below). With respect to downstream pathways, the model was restricted to ERK and Akt because, among the many cascades induced by ErbB ligands, both are known to be critical for mitogenesis and cell survival. Among the 13 known ErbB ligands, EGF and heregulin (HRG) were implemented as being representative of two major specificity classes: EGF binds only ErbB1-containing dimers, and HRG binds only ErbB3- and ErbB4-containing dimers (Berkers *et al*, 1991; Peles, 1992; Carraway *et al*, 1994). Modeling two ligands, seven receptor dimers and

downstream ERK and Akt pathways involved 28 distinct proteins. Protein–protein interactions, post-translational modifications and compartmentalization involving these 28 proteins generated a total of 471 additional species that participated in 828 reactions requiring 499 differential equations, 201 unique reaction rates and 28 non-zero initial conditions (species, rates and reactions can be found in Supplementary Figure 2). Although variations in network topology were explored in the course of model assembly, only the final configuration is described below; it represents the most compact description of the full ErbB receptor layer. Future model refinement is expected to focus on ensemble approaches that attempt to resolve topological uncertainty in the Akt pathway (Kuepfer *et al*, 2007), inclusion of additional signaling cascades and more principled treatment of combinatorial complexity during receptor activation.

On the basis of the experiments in HeLa cells, it has been proposed that EGF-induced expression of dual specificity phosphatases and other ErbB-negative regulators plays an important role in shaping the dynamics of intracellular signaling starting $\sim$30 min after ligand addition (Amit, 2007). However, when A431 or H1666 cells, which are the focus of this study, were treated with cycloheximide (an inhibitor of protein synthesis) no significant difference in ERK and Akt dynamics was observed relative to cells not treated with cycloheximide (data not shown). Thus, we ignore the impact of ligand-induced protein expression on the dynamics of immediate-early signaling in the current work.

Mass action kinetics

Reactions in the ErbB model were formulated as elementary unimolecular or bimolecular processes according to the law of mass action, with $r=kA$ for a unimolecular reaction involving protein A (where A indicates the total number of molecules of A per cell, r a rate and k a rate constant), $r=kAB$ for a bimolecular reaction involving A and B. Hill functions and other higher order algebraic expressions were not used because they represent approximations to cascades of elementary reactions. Thus, cooperativity, nonlinear input–output behaviors and feedback arise in the model only from the interplay of simple reactions. Protein concentrations throughout the ErbB network were high (absolute protein numbers $\gg 10^3$ per cell), so deterministic approaches were used. We have not yet considered the possible involvement of slow reactions or small reaction compartments ($<<100$ molecules) for which stochastic simulation might be more suitable.

Compartmentalization

Biological and reaction compartments were implemented and both were assumed to be well mixed. The former included compartments for plasma and endosomal membranes, cytosol, nucleoplasm and lysosomal lumen. We also implemented clathrin-mediated endocytosis as a second-order reaction in clathrin and ErbB; this is obviously an extreme simplification of the actual biochemistry but does reflect the need for clathrin and the receptor to interact prior to vesicular uptake. Reaction compartments were implemented by representing a single-gene product as multiple species each in its own well-mixed pool and able to participate in its own set of reactions. This made it possible to model the actions of scaffolding and adapter proteins, the molecular details of which are unclear. Protein transport was modeled in a computationally tractable manner as movement of a species from one compartment to the next with first-order kinetics (spatial gradients and partial differential equations were therefore avoided). In the current model, reaction compartments were used to encode cytosolic and membrane-bound Ras and to represent protein phosphatase 2A (PP2A), an enzyme that dephosphorylates Raf, MEK and Akt in IERMv1.0 (Ugi *et al*, 2002). Although the model includes only one PP2A catalytic subunit, by implementing pools of PP2A in three separate compartments it was possible to account for differential phosphatase activity with respect to Raf, MEK and Akt. Differential activity presumably arises from the association of PP2A with different regulatory subunits (Silverstein *et al*, 2002), but we did not include any of these details.

Simplifying network structure

The process of multi-site receptor phosphorylation and recruitment of SH2-containing proteins to receptor tails generates a 'combinatorial explosion' in possible species and reactions (Hlavacek, 2003). If we assume that the 12 known SH2-binding phospho-sites on ErbB1 are independent of each other and exist in both phosphorylated and unphosphorylated states, then ErbB1 monomers have 2^{12}=4096 distinct states and dimers have $(4096)^2 \sim 1.6 \times 10^7$ states (Schulze *et al*, 2005). Careful treatment of these phospho-states is likely to be important, as ErbB1 binds $\sim$8 distinct SH2 proteins and in some cases, multiple SH2 domains compete for binding to a single tyrosine residue (Jones *et al*, 2006). However, systematic computational treatment of combinatorial complexity places severe demands on parameter optimization because repeated cycles of numerical integration are required (Blinov *et al*, 2004). For simplicity, we therefore used a two-state approximation in which receptors were either fully dephosphorylated or fully phosphorylated. The two-state assumption seems reasonable, given experimental data that five ErbB1 phosphotyrosine sites (Y845, Y992, Y1045, Y1068 and Y1173) for which assays could be established in A431 cells were phosphorylated and dephosphorylated with similar dynamics over a 120-min period post-EGF stimulation (data not shown). However, the assumption is likely to break down under some conditions (Schulze *et al*, 2005). Thus, our approach should be considered a first step toward a more complete treatment of receptor biochemistry as computational methods advance and additional experimental data are collected.

Determining parameter values

ODE models are composed of initial-value differential equations involving two types of parameters that must be measured or estimated: initial species numbers (protein concentrations, expressed in molecules per cell) given by $x_{0,i}$ for the ith species, and rate constants (forward and reverse rate constants for complex formation and enzymatic rate constants for enzymes) given by k_j for the jth rate. Prior to estimation, we specified parameter values for as many $x_{0,i}$ and k_j as possible

Table I Classes of reactions and nominal values for their rates

Reaction type	Forward (per cell)	Reverse
EGF receptor binding[a]	1.0×10^{7}/M/s	3.0×10^{-2}/s
HRG receptor binding[a]	1.0×10^{7}/M/s	7.0×10^{-2}/s
Receptor dimerization[b]	1.6×10^{-6}/molecule/s	1.6×10^{-1}/s
Internalization[c]	1.3×10^{-3}/s	5.0×10^{-5}/s
k_{cat} phosphorylation[d]	1.0×10^{-1}/s	—
k_{cat} dephosphorylation[e]	1.0×10^{-2}–1.0×10^{-3}/s	—
Adapter binding to receptor/protein–protein complex formation[f]	1.0×10^{-5}/molecule/s	1.0×10^{-1}/s
Grb2–Sos binding[g]	7.5×10^{-6}/molecule/s	1.50/s

The reaction class was used to determine the values of rate constants at the start of calibration runs. Nominal values were obtained from previously published measurements or theoretical and computational estimates.
[a]Berkers *et al* (1991); Landgraf and Eisenberg *et al* (2000)
[b]Graus-Porta *et al* (1997); Hendriks *et al* (2003).
[c]For ErbB1 homodimers, Graus-Porta *et al* (1997); Hendriks *et al* (2003).
[d]Yun (2007).
[e]The k_{cat} of dephosphorylation was assumed to be 'slow' relative to off-rates of phosphatase and target association, consistent with the assumption that the catalytic step is rate-limiting in enzymatic reactions.
[f]Estimated with diffusion and geometric considerations; Northrup and Erickson (1992); Felder (1993); Kholodenko *et al* (1999).
[g]Sastry (1995).

based on literature data (Table I). In addition, $x_{0,i}$ was measured for several key proteins in A431, H1666 and H3255 cell lines (ErbB1–4, Shc, MEK, ERK and Akt) by semiquantitative immunoblotting relative to recombinant standards; our measurements (e.g. $\sim 10^6$ molecules of ErbB1 per A431 cell) were consistent with literature estimates when available. Receptor–ligand association constants for EGF and HRG were obtained from published cell surface-binding assays or surface plasmon resonance experiments performed on purified receptor ectodomains (Berkers *et al*, 1991; Landgraf and Eisenberg, 2000; Stein *et al*, 2002). Rate constants for ErbB1/1 and ErbB1/2 dimerization were also obtained from the literature (Hendriks *et al*, 2003). Association rates for other receptor dimers were inferred from published data describing partnering preferences for ErbB receptors (Graus-Porta *et al*, 1997). Experimentally determined rates of internalization were used for ErbB1 homodimers (Hendriks *et al*, 2003). Downregulation of ErbB3 and ErbB4 was not modeled explicitly, except for inclusion of a single tyrosine phosphatase of unspecified identity that acted on all four receptors. We observed that ErbB3 activation peaks and is substantially downregulated on the order of minutes (data not shown). Ubiquitin-mediated protein degradation has been shown to regulate ErbB3 (Cao, 2007), but this process is on the order of several hours, which is too slow to account for the observed dynamics of ErbB3 phosphorylation. Further experimental analysis of ErbB3 and ErbB4 downregulation is therefore necessary to enable detailed modeling. The remaining class of rate constants covered the binding of cytoplasmic proteins to receptor tails and to partner proteins (e.g. PI3K to Gab1, MEK to ERK). Some of these rates have been measured experimentally for ErbB- or other RTK-mediated signal-transduction systems, including the forward and reverse rates for PI3K binding to the insulin receptor (Felder, 1993), and others have been estimated computationally under the assumption that translational and rotational diffusion-limited kinetics prevail (Northrup and Erickson, 1992). On the basis of rates of related reactions and theoretical estimates, we assumed adapter binding to receptors to have nominal forward rates of 10^{-5}/molecule/s and reverse rates of 10^{-3}/s on a per cell basis, as first described by Kholodenko *et al*, (1999). However, Grb2–Sos dissociation rates were set at a value slightly lower than nominal adapter-binding off-rates to reflect the experimental observation that the two proteins are constitutively complexed (Sastry, 1995).

Parameter optimization

Parameters for which good experimental cell-based estimates are available, including K_d values for EGF receptor and HRG receptor binding and the concentrations of experimentally measured species were set at their measured values. Other parameters were estimated by minimizing an objective function (see Materials and methods) comprising the normalized root mean square deviation (RMSD) between time course data and computed model trajectories. Training data comprised 120 measurements collected in triplicate for pErbB1, pAkt and pERK levels at 10 time points from 0 to 120 min following stimulation of A431, H1666 or H3255 cells with saturating (5 nM EGF or HRG) or near-physiological (10 pM EGF or 100 pM HRG) ligand concentrations. Assays were performed with phospho-specific antibodies for ErbB1-Y1068, ERK1-T202/Y204 (ERK2-T185/Y187) and Akt-S473 as the phosphorylation states of these kinases are known to be reasonable proxies for activity, as well as for Shc-Y239/Y240, which is phosphorylated by ErbB1. A431 cells were chosen largely for their high expression of ErbB1, which ensures that receptor modification was easily measurable, and also because A431 cells have been the subject of many previous studies (Gadella and Jovin, 1995; Graus-Porta *et al*, 1997; Jones *et al*, 2006). H1666 and H3255 cell lines, which carry wild-type and mutant L858R ErbB1 receptors (Sordella *et al*, 2004; Tracy, 2004), respectively, were studied to ascertain whether the IERMv1.0 model could effectively capture differing signaling dynamics in other cell types.

On the basis of previous work with models of similar complexity (Singer *et al*, 2006), we expected IERMv1.0 to be non-identifiable (Melke, 2006; Gonzalez, 2007) and the landscape of the objective function to contain multiple local minima. This precluded parameter optimization through local optimization methods such as Levenberg–Marquardt (Marquardt, 1963) or simple measures of convergence such as Fisher information (Rodriguez-Fernandez *et al*, 2006) and instead required broad and repeated searches through parameter space for good fits. We therefore used simulated annealing (SA) to search across a region of parameter space spanning 2.5 log orders above and below *a priori* values (as described in Table I) (Kirkpatrick *et al*, 1983). By restricting ourselves to a subset of 75 rate constants and initial conditions with the greatest impact on the objective function (out of 229 total; as identified by sensitivity—see below), we substantially improved the convergence of parameter optimization. SA and similar methods have been applied widely to estimate parameters for biochemical reaction models (Gonzalez, 2007), but not, to our knowledge, with a set of ODEs as large as those in IERMv1.0. We tested the performance of SA using synthetic data comprising the same measurements and time

points as real data but generated with a model fitted manually to A431 time course. Twenty rounds of SA were observed to return excellent fits of training data to simulation ($R^2 > 0.98$; see Materials and methods and Supplementary Figure 3 for details), showing that SA is an effective means to optimize parameters in IERMv1.0. Even with synthetic data obtained from a model with known parameters, SA returned fits in different areas of parameter space that were associated with different values of the objective function. These differences arise when annealing does not converge rapidly, often because the landscape of the objective function is flat in multiple dimensions and lacks well-defined minima (Brown and Sethna, 2003), but differences that are smaller than measurement error (which we estimate at ± 10% in our experiments) are not of practical significance because they cannot be distinguished experimentally. We therefore considered all fits having an RMSD between simulated trajectories and experimental data of <0.1 as equally valid. Such fits could be recovered for actual data derived from multiple cell lines, despite substantial differences in pErbB1, pERK and pAkt dynamics (Figure 2A–C). However, SA with real data never returned the near-perfect fits observed with synthetic data (Supplementary Figure 3), particularly in the case of low HRG concentrations, suggesting the presence of errors in model structure. We also found it difficult to train the model simultaneously to Akt dynamics at high and low EGF doses: experiments showed a 50-fold increase in EGF concentration from 0.01 nM to 5 nM to have little effect on maximal Akt levels, but to induce different dynamics of downregulation (Figure 2A, A431). No fit has as yet reproduced this effect and further refinement of Akt biochemistry is probably necessary.

The number of identified good fits was constrained by computational demands; on average, finding one good fit required 100 annealing runs and 24 h on a 100-node cluster computer. As the amount of training data increased, so did the number of constraints on the objective function, increasing the

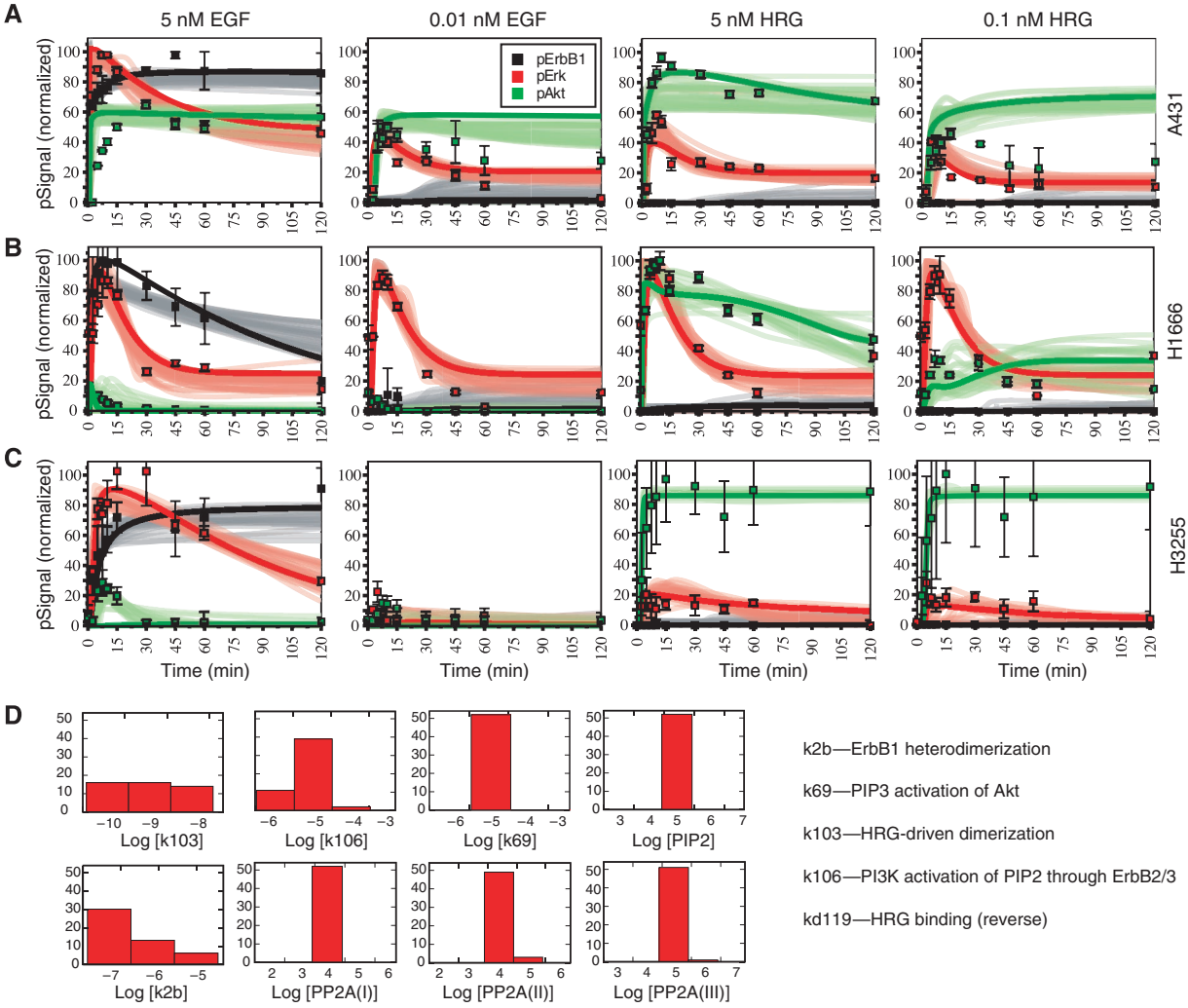


Figure 2 Experimental and simulated dynamics for three cell lines, each treated with two concentrations of EGF and HRG. Twenty good fits are shown for each cell line. (**A**) A431, (**B**) H1666 and (**C**) H3255 were stimulated with 5 nM or 0.01 nM EGF or 5 nM or 1 nM HRG as indicated. Vertical axes denote normalized signals representing the phosphorylated (and nominally active) forms of ErbB1, ERK and Akt such that the highest measured and simulated value for one signal, across all treatment conditions, was set to 100. In each panel, the fit having the lowest value of the objective function is indicated by a bold line; slightly less good fits are indicated by faded lines. Experimental data are indicated by dots and standard error of the mean (biological triplicates) by the vertical bars. (**D**) Distributions of 8 parameter values across 50 independent parameter optimization runs; parameter identities are described in more detail in Supplementary information.

ruggedness of the landscape and the time needed to find minima. However, the challenge of parameter estimation using time course data is more fundamental than this. Sethna and others have shown the process to be 'sloppy' so that even 'ideal and complete' synthetic training data allow only a subset of the parameters in a complex biochemical model to be estimated (Brown and Sethna, 2003). Such parametric uncertainty is a reality for all complex biochemical models and remains an issue even when data are added and different fitting methods are implemented.

When different rounds of SA having similar final goodness of fit to data (that is, RMSD <0.1) were compared, some parameter values lay within $\sim$3-fold of their mean values, whereas others took on the full range of values allowed in the search (Figure 2D; Supplementary Figure 4). This confirms our expectation that IERMv1.0 is non-identifiable. Parameters with similar values across multiple fits are well constrained, whereas those with widely varying values are poorly constrained, but we have not yet recovered enough fits (at a cost of $\sim$2000 processor hours per fit) to compute uncertainties for individual parameters. Future implementation of refined sampling approaches such as Bayesian calibration and Markov chain Monte Carlo methods will be required (Vyshemirsky and Girolami, 2008).

Analyzing partially calibrated models

To ascertain whether biologically meaningful information can be derived from the IERMv1.0 model despite its non-identifiability, we examined five good fits for features that seemed to be well constrained based on parameter sampling. For the sake of simplicity, we restrict this discussion to data obtained from A431 cells, but similar results were obtained for other cell lines. First, we sought to identify, by computing normalized sensitivities, those parameters having the greatest impact on biological outputs. The fully normalized sensitivity ($s_{ij}(t)$) of the ith observable $c_i(t)$ (such as pERK levels) with respect to a change in the jth rate constant (k_j) or initial concentration ($x_{0,j}$) is given by:

$$s_{ij}(t) \equiv \frac{\partial \ln(c_i(t))}{\partial \ln(k_j)} \tag{1a}$$

$$s_{ij}(t) \equiv \frac{\partial \ln(c_i(t))}{\partial \ln(x_{0,j})} \tag{1b}$$

Sensitivities were converted into time-independent quantities by integration:

$$s_{ij} \equiv \frac{1}{T}\int_0^T \mathrm{dt}\cdot|s_{ij}(t)| \tag{2}$$

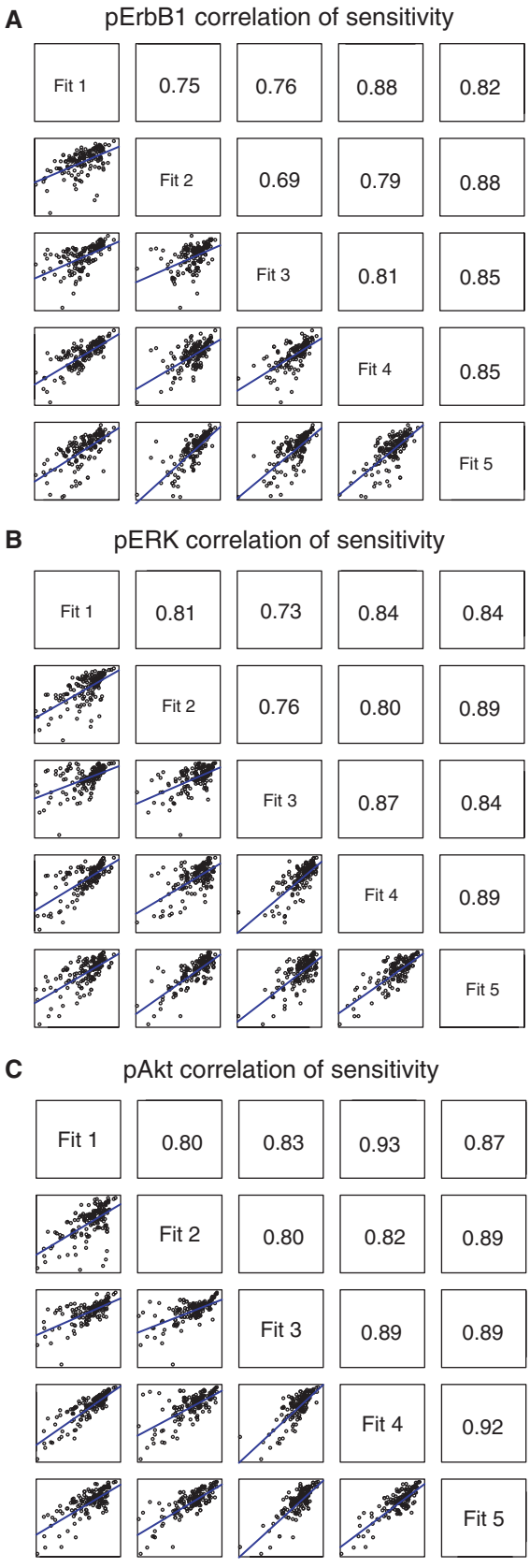
where T is the final time point, and the absolute value of the integrand ensures that negative and positive sensitivities do not trivially cancel to zero under the integral. The quantity s_{ij} measures the fractional change in the ith observable upon a fractional change in the value of the jth parameter, normalized by $1/T$ so as to obtain a time-averaged value. Outputs of interest ($c_i(t)$) include the trajectories of phosphorylated ErbB1, ERK and Akt over time. These were, in general, aggregates of multiple model species. For example, total pERK levels are represented in IERMv1.0 by $\sim$30 differentially bound and localized species. When local sensitivity analysis for pErbB1, pERK and pAkt was performed on the five best fits to data from A431 cells, s_{ij} ranged from 0 to $\sim$0.8, depending on the parameter. By plotting all pairs of s_{ij} values for all pairs of fits, correlations of R=0.69 to 0.93 were observed (where R is the correlation coefficient; Figure 3). Correlation of less than 1.0 is expected, as sensitivity is a local property dependent on actual position in parameter space, which varies from fit to fit, but the mean value of R=0.84 implies significant similarity in sensitive parameters across fits. We explored this issue further using a more robust but less familiar measure of 'regional' sensitivity:

$$s_{ij}^{\mathrm{Regional}} \equiv \left\langle \frac{1}{\tau_2-\tau_1}\int_{\tau_1}^{\tau_2} \mathrm{d}t|s_{ij}(t)| \right\rangle_{\mathrm{AVG}} \tag{3}$$

where $\langle \ldots \rangle_{\mathrm{AVG}}$ indicates an average over parameter values lying within $10^{2.5}$-fold above and below the parameter values of the fit. This method of calculating average local sensitivity has been discussed previously (Cukier *et al*, 1973; Bentele, 2004; Mahdavi, 2007) and has two advantages over simple sensitivity: (i) it accounts for the possibility that a fit does not precisely hit a nearby minimum due to problems with convergence during SA and (ii) it reduces the impact on sensitivity of small but steep irregularities in the landscape of the objective function. Because averaging over the entire region of parameter space was too costly computationally, we approximated the average by randomly sampling for parameter space in the vicinity of the fit. Sampling revealed a power-law relationship that reached convergence at $\sim$1000 rounds (Figure 4A); we therefore performed regional sampling at 1000 points around the nominal fit. Regional sensitivity of five A431 model fits showed substantial commonality in the rank order: 8 of the 10 most sensitive parameters for EGF- or HRG-stimulated pERK were shared across five fits as were 7 of 10 sensitive parameters for pAkt dynamics across three models (Figure 4). Thus, even a partially calibrated and non-identifiable model yields meaningful information on parameter sensitivities, probably because the most sensitive parameters are among the best constrained. In the following, we concentrated on a single good fit to A431 data that had the lowest value of the objective function after calibration.

Time, stimulus and signal dependence of sensitivity analysis

Sensitivities are always calculated for a particular model feature (e.g. time-integrated pERK or pAkt values) under a particular condition (e.g. stimulation with EGF or HRG). Correlation plots were used to determine how the identities and rank order of parameter sensitivities depended on the feature and condition (Figure 5). Most s_{ij} fell close to the origin, demonstrating that only a few parameters impacted each feature, but sensitive parameters exhibited significant differences from one feature to the next. For example, sensitive parameters for pERK activation by EGF or HRG stimulation were largely shared (Figure 5A). In contrast, when factors



controlling Akt activation by EGF and HRG were compared, ~50% of the sensitive parameters lay well off the diagonal, demonstrating independent control of Akt by the two ligands (Figure 5B). Clustering of sensitive parameters away from the diagonal was even more pronounced when pERK and pAkt were compared (Figure 5C and D). Finally, differences in the identities of highly sensitive parameters were observed for high and low ligand concentrations and early and late time points (Supplementary Figure 5). To confirm these findings experimentally, we selected a non-obvious differential sensitivity for analysis. Modeling suggested that pAkt levels were more sensitive to changes in ErbB1 activity than those of pERK (Figure 5C). We therefore measured the dose-dependent inhibition of pERK and pAkt, in EGF-stimulated A431 cells by the ErbB1-specific inhibitor gefitnib and the dual-specific ErbB1–ErbB2 inhibitor lapatinib. When sensitivity to ErbB1 activity was defined as the concentration of drug that inhibited the downstream signal by 50% (IC_{50}), we observed Akt phosphorylation to be significantly more sensitive to gefitinib ($IC_{50} \sim 5 \times 10^{-8}$ M) than ERK phosphorylation ($IC_{50} \sim 5 \times 10^{-7}$ M) (Figure 6A). A similar 10-fold difference was observed with lapatinib (Figure 6B), confirming our prediction from sensitivity analysis. We therefore conclude that parameter sensitivity in the ErbB model is critically dependent on the choice of target observable $c_i(t)$.

The target and context dependence of parameter sensitivity is fully consistent with elementary dynamical systems theory, but our work shows that it is also true for physiologically important outputs and realistic parameter sets in biochemical models. This fact is often ignored in discussions of robustness, in which it is claimed that parametric insensitivity is inherently physiologically meaningful (Morohashi, 2002). Because the parameters sensitive for different observables or conditions overlap only partly, the fraction of parameters that are sensitive (42 out of 229) is significantly greater than what might be assumed from examining only one observable or condition (for which 6–10 sensitive parameters is typical). For example, although pAkt dynamics in EGF-stimulated cells are insensitive to changes in the initial concentration of the GTPase Ras (Figure 5B), this does not mean that the ErbB pathway is robust to changes in Ras levels: pERK dynamics are critically dependent on this parameter (Figure 5A, yellow dots).

With respect to the utility of sensitivity analysis on a partially constrained model, we note that considerable biological insight can be obtained by comparing sensitive parameters for different $c_i(t)$. For pERK dynamics in EGF- or HRG-stimulated cells, the three sets of sensitive parameters were Ras levels, on- and off-rate constants for association of MEK and Raf, and for association of MAPK pathway phosphatases and their targets (Figure 5A). This makes sense biologically because both EGF and HRG activate ERK through Ras and Raf. In contrast, the differential sensitivity of pERK to

Figure 3 Correlation of sensitivities between five best A431 fits that have been optimized for parameters by simulated annealing. Panels give time-integrated sensitivities for (**A**) pErbB1, (**B**) pERK and (**C**) pAkt dynamics. Within each panel, small boxes in the lower triangle show the correlation plot for parameters sensitivities between two fits; upper triangle shows correlation value (R) associated with each correlation plot.

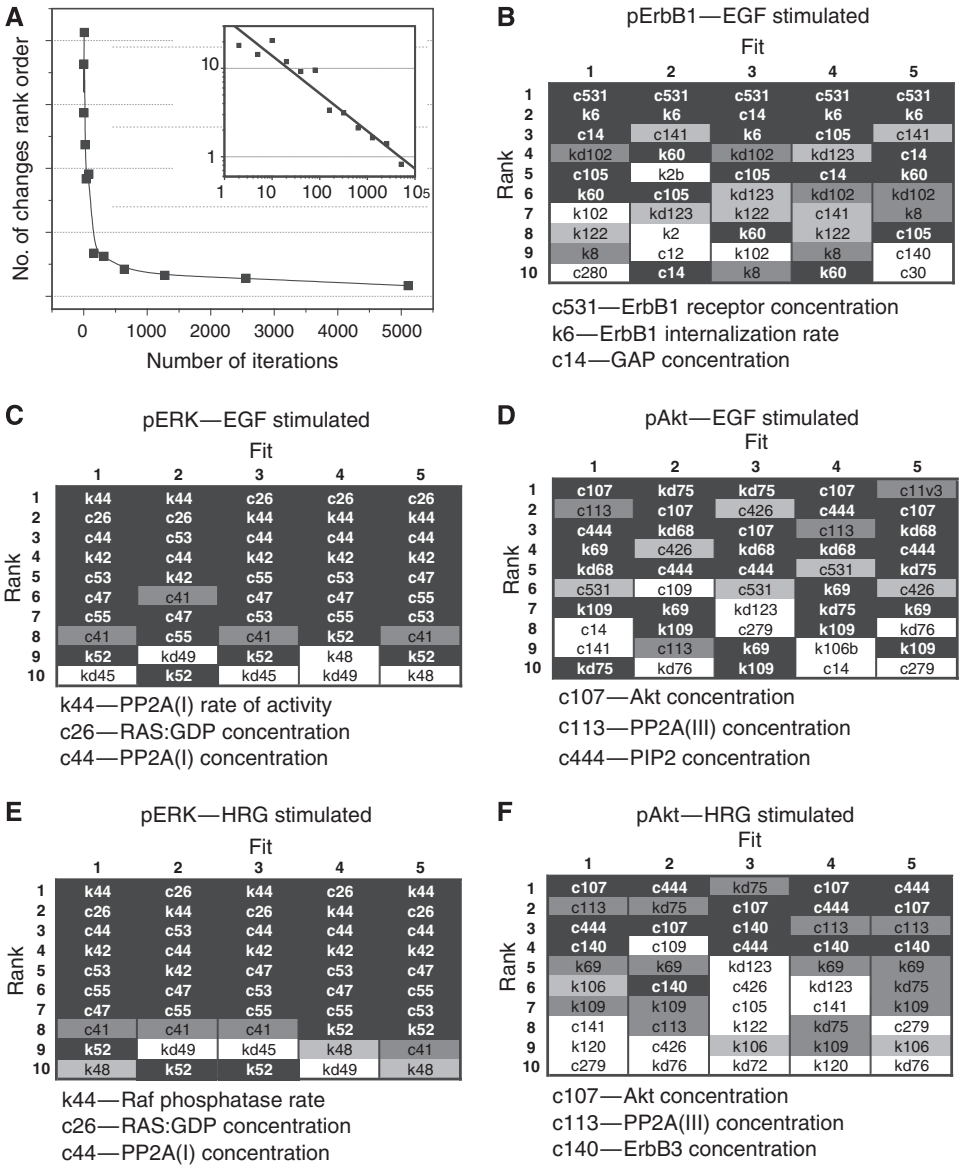


Figure 4 Regional sensitivity analysis. Regional sensitivities were calculated according to equation (3). (**A**) Changes in sensitivity rankings as a function of number of randomization iterations used in sampling for regional sensitivity. Changes in rankings were calculated relative to the rankings of the run with the next smallest number of iterations. (Inset) Power-law relationship between rank changes and increasing sample size on a log-linear scale with an exponent of 0.4. (**B**–**F**) Regional sensitivities of five fits to A431 data, showing the top 10 sensitive parameters for pErbB1 (B), pERK (C, E) and pAkt (D, F) dynamics integrated over time following stimulation with 5 nM EGF (B–D), or HRG (E, F). Shading of entries is as follows: black—parameters with a rank order in the top 10 across all five fits; dark grey—in top 10 across four fits; light grey—in top 10 across three fits.

ErbB1 initial concentrations in EGF-stimulated cells reflects the key role played by ErbB1/1 dimers in EGF binding, whereas the sensitivity of pERK in HRG-stimulated cells to ErbB2 and ErbB3 levels reflects their roles in HRG binding. The number of ligand-specific sensitive parameters for pAkt dynamics was particularly high, revealing non-overlapping mechanisms of Akt activation by EGF and HRG. For example, for HRG but not EGF stimulation, on-rates of binding of PI3K–p85 to ErbB3 receptor were critical. Intriguingly, PIP2 levels, and rate constants for PIP3–Akt and PIP3–PTEN interaction appeared to play particularly important roles in pAkt dynamics regardless of ligand, consistent with frequent mutation of PTEN and regulators of inositol phosphates in human cancer.

The set of sensitive parameters common to both EGF-mediated pERK and pAkt induction was limited to GAP and ErbB1 initial concentrations (Figure 5C). This is reasonable given that both ERK and Akt depend on ErbB1 and Ras for activation. In contrast, protein phosphatases appeared as significant off-diagonal factors, implying differential control of the two kinases. The calibrated v_{max} (calculated using the expression $v_{max}=k_{cat}A$, where A is the concentration of enzyme A) for the PP2A compartment that targets pRaf and pMEK was 1–3 orders of magnitude greater than for the compartment targeting pAkt (data not shown). Differential sensitivity extended to all three PP2A reaction compartments, supporting the hypothesis that PP2A functions quite differ-

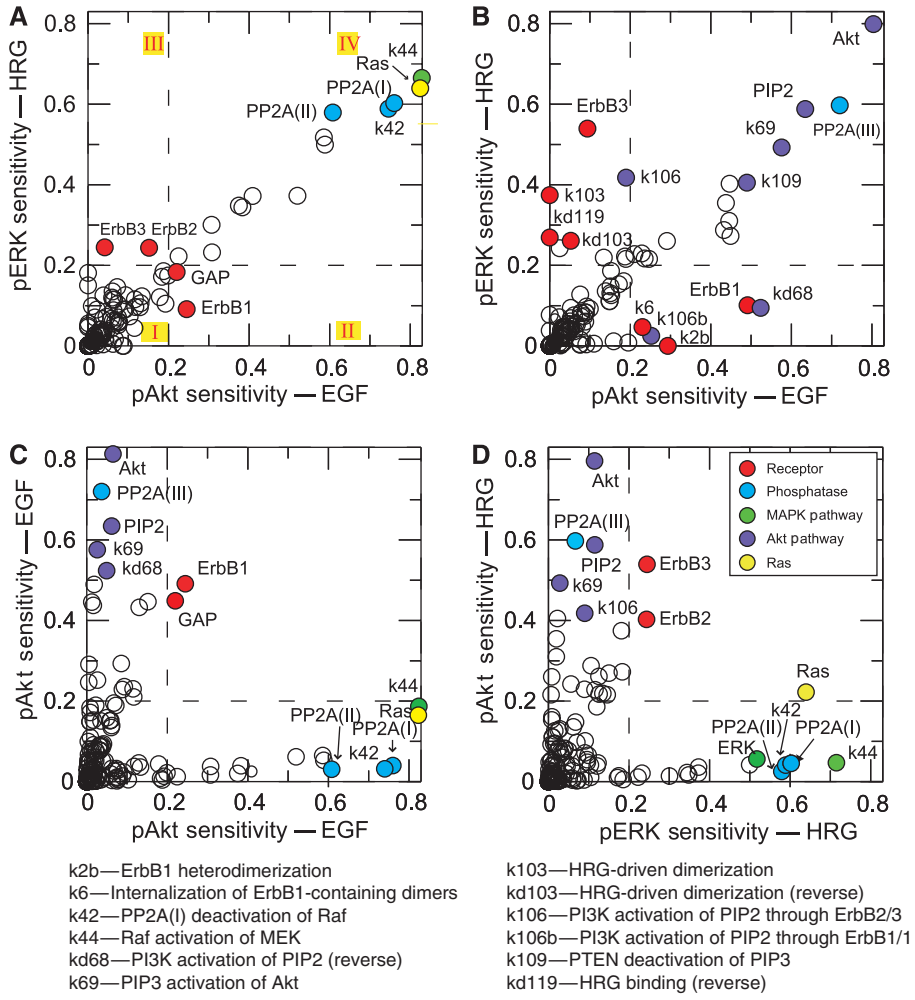


Figure 5 Dependency of parametric sensitivity on model feature or condition. Sensitivities were calculated according to equation (1). (**A**–**D**) Correlation plots comparing two dynamic variables (model features) or ligands (conditions). Vertical and horizontal dotted lines divide the plots into four quadrants (as labeled in red on a yellow background): I—insensitive parameters; II–III—parameters exhibiting differential sensitivities; IV—parameters along the diagonal showing joint sensitivities. Colors depict the biological activity or part of the pathway to which a parameter belongs as follows: red—receptor; blue—phosphatase; green—MAP kinase cascade; purple—PI3K–Akt cascade; yellow—Ras and its regulators. (A) pERK sensitivities compared for stimulation with 5 nM EGF versus HRG stimulation and (B) pAkt sensitivities under the same conditions. pERK versus pAkt sensitivities compared under 5 nM EGF (C) or HRG (D).

ently on pERK and pAkt. As a computational test of this prediction, we implemented an alternative model having only one compartment for PP2A. In this model, pERK dynamics deviated substantially from experimental values under conditions of HRG stimulation (Supplementary Figure 6, dotted line). However, such a comparison ignores the possibility that a single-compartment model with an optimized set of parameters might perform as well as the three-compartment model. When we refitted the single-compartment model using data from A431 cells and SA, we observed substantial improvement in goodness of fit. Nonetheless, significant differences between single- and three-compartment models were obtained across multiple SA runs: IERMv1.0 correctly predicted that pERK levels would peak at 5 min after HRG stimulation before gradually declining (Supplementary Figure 6, solid line), whereas no single-compartment model returned a fit with correct timing (typically, the pERK peak was delayed ~15 min in simulation relative to experiment; Supplementary Figure 6, dashed line). The inability of SA to identify good fits to data for the single-compartment model does not constitute proof that PP2A exists in multiple distinct states, but the negative result is consistent with the hypotheses that dephosphorylation of Raf, MEK and Akt occurs at different rates, presumably through different PP2A-containing complexes. Moreover, our findings are consistent with previous analysis suggesting that phosphatases play a particularly critical role in shaping RTK signaling dynamics (Heinrich *et al*, 2002). From a methodological perspective, our findings also highlight the need to re-optimize parameters when comparing models that differ in structure.

Network-level predictions and their experimental confirmation

An important measure of the value of a kinetic model is its ability to generate testable predictions that provide new insight into complex biochemical processes. Experimental validation

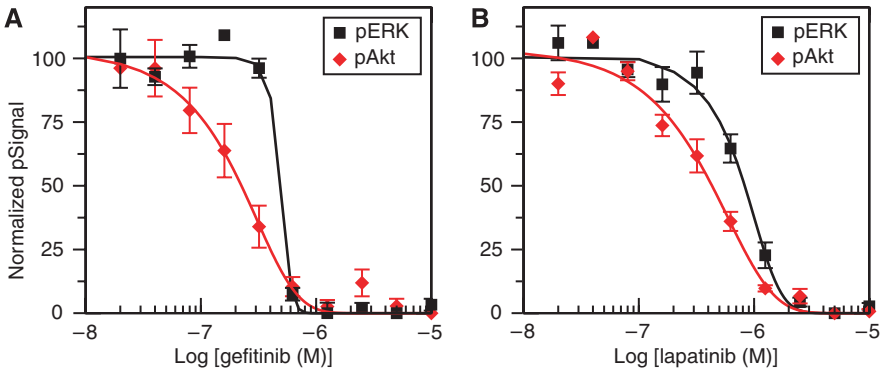


Figure 6 Differential sensitivity of pERK and pAkt to gefitnib treatment (**A**) and lapatinib treatment (**B**). Experimentally measured pERK (red diamonds and lines) and pAkt (black squares and lines) 5 min after EGF stimulation in A431 cells treated with inhibitor 10 min prior to ligand addition. Error bars indicate standard error of the mean.

of pAkt sensitivity to ErbB1 inhibition relative to pERK is one example (Figure 6). The trained IERMv1.0 model also predicted correctly the dose–response of proteins such as pShc (Figure 7A), but this is not a stringent test of performance as pShc levels closely followed those of pErbB1. We therefore turned to predicting and understanding the dose–response behavior of the ErbB network, a fundamental measure of biological activity that has been examined in detail by others (Goldbeter and Koshland, 1984; Huang and Ferrell, 1996; Ferrell and Machleder, 1998; Hendriks *et al*, 2005) Two measures of dose responsiveness are the ligand concentration for half maximal activation (the EC_{50}) and the steepness of the transition from on to off, which can be described by an apparent Hill coefficient (H_{app}). H_{app} is calculated by a nonlinear least-squares fit of dose–response data to the Hill equation:

$$\mathrm{pSignal}(x) = \frac{x^{H_{app}}}{x^{H_{app}} + k^{H_{app}}} \quad (4)$$

where pSignal is the normalized magnitude of signal (e.g. ErbB1 phosphorylation), x is the concentration of input (e.g. EGF), the free parameters H_{app}, k are the apparent Hill coefficient and half-maximal concentration of x. A sigmoidal dose–response curve for a second-order ligand–receptor interaction has $k{=}K_d$, $H_{app}{=}1.0$ such that the levels of receptor–ligand complex rise from 10 to 90% of their maximal value over an 81-fold change in ligand concentration. A third useful measure of a signaling cascade is the extent of amplification and attenuation at successive steps:

$$\Psi(\mathrm{EGF}, t) = \frac{\mathrm{pSignal}_1(t)}{\mathrm{pSignal}_2(t)} \quad (5)$$

where $\mathrm{pSignal}_n$ refers to the activity or state of modification of protein n in a signaling cascade. Because all signals are dynamic, EC_{50} and H_{app} are functions of time, which we set at $t{=}5$ min (the peak for transient ERK and Akt responses).

Simulation across multiple fits showed H_{app} for pERK and pAkt as a function of EGF (equation (4)) to be $\sim$0.30 and, unexpectedly, that both kinases would be activated to significant levels in A431 cells at ligand concentrations well below the lowest reported K_d for EGF–ErbB1 association ($\sim$0.1 nM; Carraway *et al*, 1994). Experiments confirmed this prediction: although pErbB1 and pShc levels rose to measurable levels only above 10^{-9} M EGF, pERK and pAKT were $\sim$20% maximal at EGF concentrations as low as $\sim 10^{-12}$ M (Figure 7A). Thus, the relationship between ligand concentration and ERK and Akt activities was nearly log-linear over a 10^6 range of ligand (between 10^{-14} and 10^{-8} M EGF; $R^2 \sim 0.98$ for a log-linear fit). Modeling suggests that this linearity arises because the ErbB signaling cascade downstream of pShc acts as a dose-dependent amplifier with a peak at $\sim 10^{-12}$ M EGF and Ψ (EGF, $t{=}5$ min) varying between 0 and 30 over the range 10^{-12}–10^{-6} M EGF. The prediction of dose-dependence in Ψ was confirmed experimentally (Figure 7B; Supplementary Figure 7B and C). These observations contrast with experimental and theoretical analysis of MAPK signaling in progesterone-stimulated *Xenopus* oocytes by Huang and Ferrell (1996) showing dose–response curves for MEK and ERK activation to be nearly switch-like with $H_{app}{=}1.7$ for MEK and $H_{app}{=}4.9$ for ERK. In EGF signaling, we observe exactly the opposite behavior: an input–output curve that was much more gradual than expected for normal second-order ligand–receptor binding.

To identify determinants of dose responsiveness, we examined the input–output behavior of a MAPK module isolated from IERMv1.0. Although the best IERMv1.0 fit predicted a shallow dose–response curve consistent with experimental data (Figure 8A), the behavior of the isolated MAPK module was consistent with results from Huang and Ferrell in exhibiting a steep dependence of pMEK and pERK on Raf activity ($H_{app}{=}4.6$ and 6.3, respectively) (Figure 8B). This difference can be rationalized by the fact that in the full model (and presumably also in cells) the protein cascade leading to ERK and MEK activation is driven by a Ras-dependent input, the strength of which varies nonlinearly with time and ligand dose and the cascade is therefore not at equilibrium. In the isolated MAP kinase module, ERK and MEK are activated by a Ras input of constant magnitude. This observation underscores the importance of considering signaling modules in the context of the larger networks in which they are embedded. We therefore undertook a general search for variables that might control pERK dose responsiveness in IERMv1.0. Starting with an A431 model fit having $H_{app}{=}2.9$ (for pERK), we used Monte Carlo sampling to explore single moves, or a succession of $\sim 10^5$ moves across a landscape of the 75 sensitive parameters optimized during model fitting. Among all possible single moves, none decreased H_{app} by greater than 10%, but searches

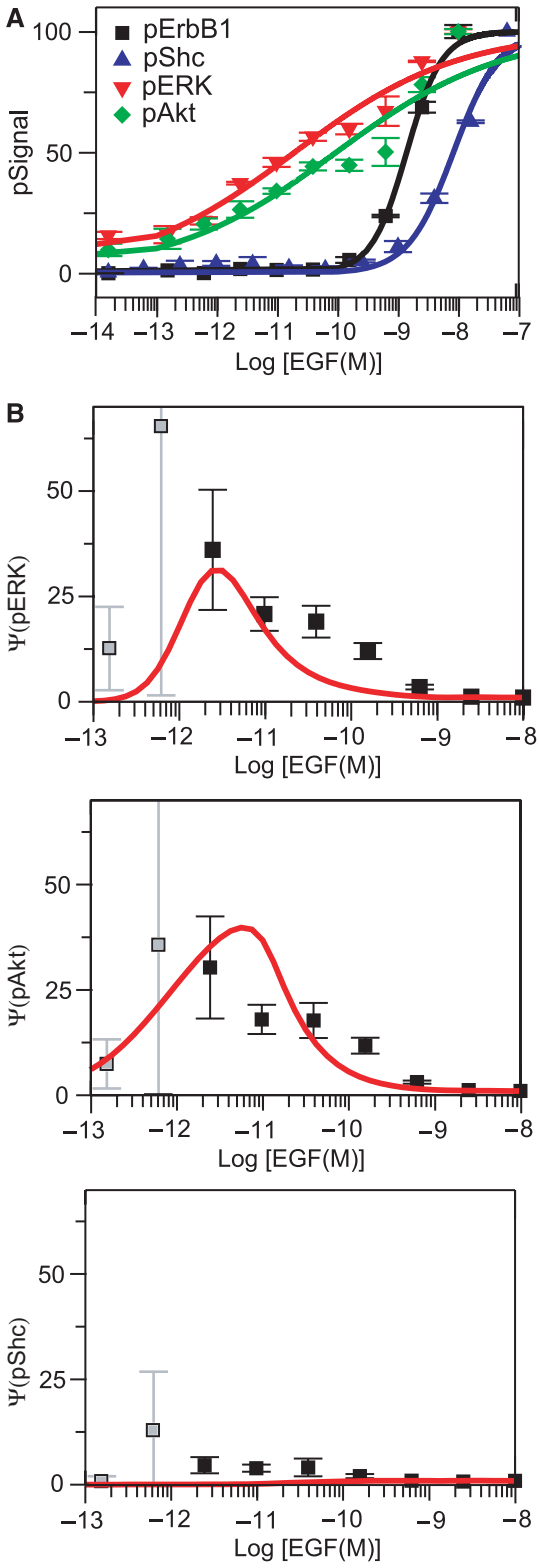


Figure 7 (**A**) Dose–response of pErbB1, pShc, pERK and pAkt (5 min after ligand addition) over a 10^6 range of EGF concentrations in A431 cells. (**B**) Amplification Ψ of signal calculated with equation (5) (5 min after ligand addition) over a 10^6 range of EGF concentrations in A431 cells for pERK, pAkt and pShc. Points represent data and lines represent simulation; error bars denote standard error of the mean. Gray boxes show high error measurements at low EGF concentrations, which arise due to phospho-ErbB1 signals being at the limit of detection.

involving a succession of changes were repeatedly successful in reducing H_{app} by 1.5- to 2-fold. Remarkably, in any Monte Carlo sampling run involving 10^5 moves, changes of only a few parameters were responsible for almost all of the change in H_{app}. This implies initial diffusive movement across a region of parameter space having nearly constant H_{app} followed by a sudden transition to a new region of space having a substantially lower H_{app} value (Figure 8C). The amount of Ras (c26), Raf–MEK association rate (k44) and MEK phosphatase rate (c53) were the parameters, the alteration of which was most frequently responsible for transitions from high to low values of H_{app} (Figure 8D). Moreover, when the impact of changes in c26 in IERMv1.0 were examined in greater detail, high Ras concentrations (10^5 molecules per cell) resulted in a steeper dose–response curve ($H_{app} \sim 1.77$), whereas low Ras concentrations (10^3 molecules per cell) resulted in a flatter curve ($H_{app} \sim 0.96$) and a higher EC_{50} for EGF (Figure 8E). Importantly, when isolated MAPK modules having high and low c26 were compared, H_{app} of ERK was unchanged at ~ 4 in both (Figure 8F). Thus, the impact of c26 on MAP signaling is restricted to settings in which the module is embedded in a larger signaling network. Similar results were obtained with Raf–MEK association rates (k44) (Figure 8G and H). From these data, we conclude that the input–output behavior of the MAPK cascade is strongly dependent on the context in which it is found: varying c26 and k44 in the context of the full IERMv1.0 changes EC_{50} and H_{app}, but varying the same parameters in an isolated cascade has virtually no effect.

Parametric sensitivity in input–output responses is something we might anticipate from the properties of electronic filters and amplifiers, but it has not previously been explored in cell signaling. The preceding analysis implies, however, that we should observe considerable diversity from one cell type to the next with respect to the EC_{50} and H_{app} for Akt and ERK. In four cancer cell lines (AU565, HBL100, MCF7 and T47D), pErbB1 activation exhibited a narrow range of EC_{50} values (10^{-8}–10^{-9} M EGF) and $H_{app} \sim 1.0$. In contrast, H_{app} for pERK and pAkt was much more variable, ranging from 0.58 to 1.61 (Figure 9A) and EC_{50} values were spread over 10^{-9}–10^{-12} M EGF (Figure 9B). These data are consistent with the biochemistry of the ErbB signaling network. Activation of ErbB1, which is most upstream and governed only by the kinetics and thermodynamics of EGF–receptor interaction and the biochemistry of the kinase domain, is expected to be similar across cell lines. In contrast, signals further down the cascade are modulated by many upstream proteins, many of whose concentrations and rate constants impact EC_{50} and H_{app}. We therefore conclude that the diversity of dose–responses that can be generated by simulation (using different sets of parameter values) also occurs in real cells. We speculate that this behavior confers great adaptability on ErbB signaling and helps to explain its importance in many aspects of cell physiology.

Discussion

In this paper, we describe the construction of an ODE-based model of immediate-early signaling involving all four ErbB receptors, analysis of parameter sensitivity and uncertainty and exploration of factors controlling overall input–output behavior. The IERMv1.0 represents ErbB receptors and down-

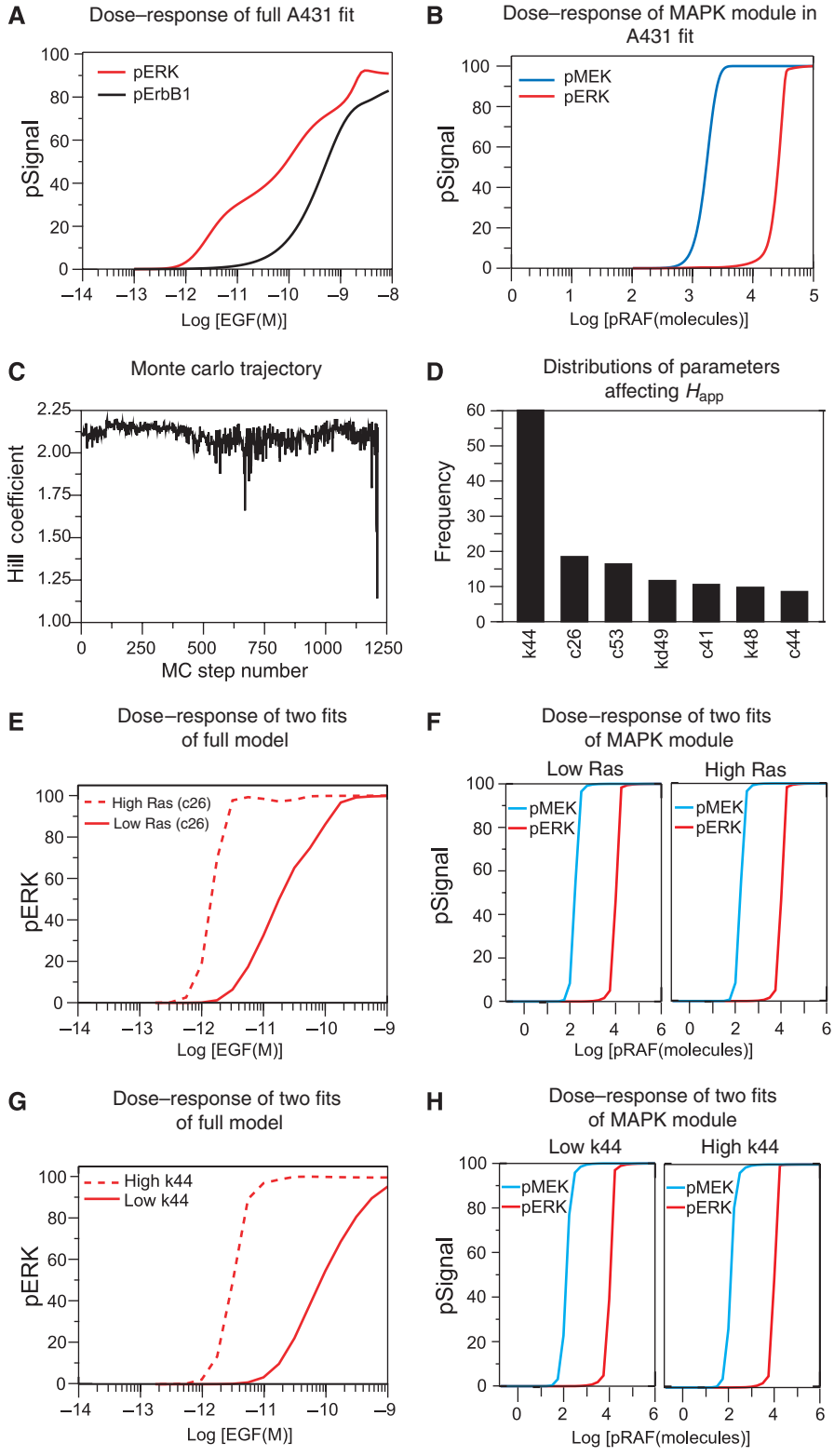


Figure 8 (**A**) Dose–response of pERK (red line) and pErbB1 (black line) computed with best IERMv1.0 fit to A431 data. (**B**) pERK (red line) and pMEK (blue line) responses in a MAPK module isolated from IERMv1.0. (**C**) Apparent Hill coefficient H_{app} (as defined using equation (4) in text) of pERK during Monte Carlo sampling of parameters. (**D**) Parameters controlling apparent Hill coefficient: c44—Raf phosphatase abundance; k48—phosphatase binding to MEK; c41—Raf abundance; kd49—phosphatase catalysis of MEK; c53—MEK phosphatase abundance; c26—Ras abundance; k44—Raf binding to MEK. (**E**) Dose–response of pERK for an IERMv1.0 fit with high Ras (parameter c26) abundance (dotted line) or low Ras (solid line). (**F**) Dose–response of pMEK (blue line) and pERK (red line) in a MAPK module isolated from a IERMv1.0 fit having low values of parameter c26 (left panel) or high values (right panel). (**G**) Dose–response of pERK for an IERMv1.0 fit with high k44 (dotted line) or low k44 (solid line). (**H**) Dose–response of pMEK (blue line) and pERK (red line) in a MAPK module isolated from an IERMv1.0 fit having low values of k44 (left panel) or high values (right panel).

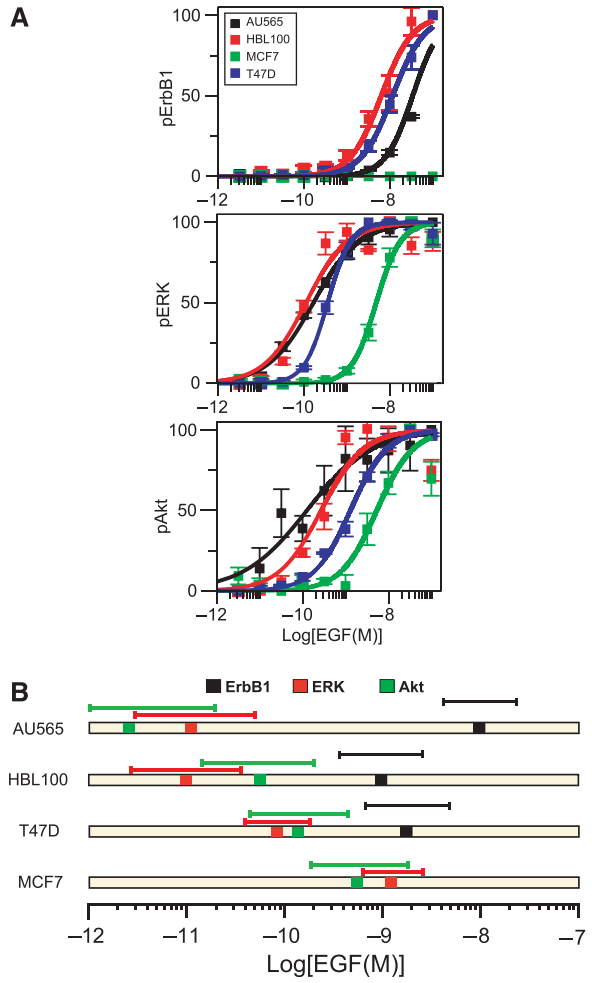


Figure 9 (**A**) Experimentally determined dose–response of pErbB1 (upper panel), pERK (central panel) and pAkt (lower panel) at t=5 min across a range of EGF concentrations in multiple cancer cell lines. Hill coefficients for each EC_{50} as calculated by nonlinear regression are given in the legend. (**B**) EC_{10}—effective concentration at 10% of maximum—of three signals for each cell line. Boxes indicate the position of the EC10. Data points and fits to a Hill curve (lines) are colored as follows: black—AU565; red—HBL100; green—MCF7; blue—T47D. The levels of pErbB1 in MCF7 were below the level of detection, but activation of ERK and Akt signals in response to EGF treatment was easily measured. Hill coefficients as calculated by nonlinear regression for activation of pErbB1, pERK and pAkt are, respectively, in AU565 cells, 1.36, 0.82 and 0.58; HBL100 cells, 1.20, 0.89 and 0.88; T47D cells, 1.18, 1.61 and 1.02. For MCF7 cells, Hill coefficients for activation of pERK and pAkt were, respectively, 1.53 and 1.03. (B) EC_{10}—effective concentration for 10% maximal activation for three phospho-signals in each cell line (box). Colored lines above boxes indicate the width of the corresponding activation profile of the Hill curve, defined as the concentration range that spans a rise in activation from 10 to 90%.

stream pathways at an intermediate level of detail in which interactions among elementary species are encoded in a continuum mass action approximation by first- and second-order elementary reactions. Our model has the merit of explicitly capturing interactions among multiple ErbB receptors and ligands, as well as downstream events such as Ras activation and MAPK and PI3K–Akt signaling, without use of complex functions, quasi-static state Michaelis–Menten approximations and aggregation of biochemically distinct species. Complex behaviors arise in the model from interactions among simple elementary reactions, making it possible to explore (through sensitivity analysis) the functions of individual species on input–output behavior and to predict the effects of RNAi, small-molecule drugs and mutations. We explore the first topic in this paper and the others in a forthcoming paper on the determinants of cellular sensitivity to anti-ErbB1 drugs (Jasper, in preparation). For simplicity, the current model omits several signaling pathways activated by ErbB receptors and lacks a rigorous treatment of the web of reaction intermediates arising from protein assembly (e.g. SH2-containing proteins). In future work, it should be possible to extend IERMv1.0 to include additional downstream signaling pathways and upstream ligands. However, thorough treatment of interactions among 13 ligands, 4 ErbB receptors and 8–12 adapters will require both conceptual and technical advances in modeling reaction networks (Hlavacek, 2003; Danos *et al*, 2007).

Our representation of immediate-early ErbB signaling involves a large number of free parameters, but with respect to scope and level of detail IERMv1.0 is not atypical of other recent efforts to model cell signaling (Kholodenko *et al*, 1999; Hatakeyama, 2003; Hendriks, 2005). Sensitivity analysis shows that the behavior of IERMv1.0 is dependent on the values of many rate constants and initial protein concentrations that have not been measured directly. Moreover, even when *in vitro* biochemical data are available, their relevance to the crowded environment of a cell is unclear (Schnell and Turner, 2004). Thus, many significant parameters must be estimated by fitting; biochemical data on isolated proteins are most useful, in this context, as a means to initiate or constrain the search (e.g. Table I). Over a series of ~2000 independent parameter optimization runs involving searches across $\sim 10^6$ parameter sets, we found only ~100 (0.01%) that had a good fit to data (RMSD < 0.1), implying that the search landscape is very rugged and deep minima are infrequent, Moreover, among equally good fits (those that varied by less than experimental error), only a subset of parameter values were constant, whereas others varied over the entire allowable range. Thus, IERMv1.0 is non-identifiable and we presume that large regions of parameter space are consistent with experimentally observed dynamics. One interpretation of this finding is that biological networks are so robust to changes in rate constants and protein concentrations (von Dassow *et al*, 2000) that parametric uncertainty is not a significant issue: key physiological behaviors, in this view, are determined primarily by pathway structure. The other extreme view is that complex models with unconstrained parameters have little practical value, as any behavior can be achieved with suitable adjustment of parameters. We have demonstrated that neither of these extremes is true: physiologically important aspects of ErbB signaling such as the dose responsiveness are indeed determined by the values of specific parameters, but useful and accurate predictions can be made using partially constrained models.

Model identifiability and parametric uncertainty

Model identifiability is a subtle issue for which a wide variety of analytical and numerical methods have been developed.

Nonetheless, most kinetic models of biological pathways are presented as though they have a single parameter solution set. In the area of biochemical or pathway models, published identifiability analysis has focused on analytic approaches to small idealized models (Margaria *et al*, 2001; White, 2001). Parameter estimation for large models such as IERMv1.0 involves numerical and Monte Carlo approaches such as SA (van Riel, 2006; Gonzalez, 2007). In general, parameter estimation has four possible outcomes (White, 2001): (i) a single, unique, parameter solution set, making the model fully identifiable, (ii) a countable number of parameter sets, making the model 'locally identifiable', (iii) an infinite number of solution sets, making the model non-identifiable and (iv) no solution sets, in which case the model probably has structural flaws. Parameter optimization with IERMv1.0 conforms to possibility (iii).

Non-identifiability of our model arises in part because the quantity, quality and type of data are inadequate and relationships among parameters and experimental observables are inherently non-unique (Hengl *et al*, 2007). Why then, do we not simply double or triple the amount of data used for model training? Unfortunately, it has been shown that even a complete set of time-resolved measurements of all species in a complex reaction model is insufficient to fully constrain all parameters due to the 'sloppy' nature of the fitting procedure (Gutenkunst, 2007). It is likely that experiments can be designed to mitigate some of this sloppiness, but the necessary design and fitting methods have not yet been developed. Thus, non-identifiability is likely to remain an issue for all complex biochemical models. Methods such as ensemble modeling and parameter sampling have been developed to cope with non-identifiability in fields as diverse as climate and nuclear reaction modeling (Christie *et al*, 2005). Means to sample parameters have also been applied to models of *Drosophila melanogaster* segment polarity (Ingolia, 2004) and mTOR signaling (Kuepfer *et al*, 2007). However, the issue of identifiability is widely ignored in biological modeling. It would be valuable to know which parameters are well constrained, which co-vary and which are truly unconstrained. The width of the distribution in parameter values across repeated rounds of SA serves as a proxy for this parametric uncertainty (Figure 5E), but the small number of good fits available for the ErbB model prevents our deriving reliable estimates for individual parameters. Nonetheless, it is clear that some parameters are much better constrained than others (e.g. phosphatase levels were tightly distributed across multiple parameter optimization runs but many receptor dimerization rates assumed a wide range of values). In principle, it is legitimate to draw conclusions based on well-constrained parameters despite the presence of other unconstrained variables. We approximate this ideal by comparing features across multiple fits. We anticipate that future implementation of SA, genetic and hybrid algorithms that take advantage of modularity or branch-and-bound (Singer *et al*, 2006) will improve our understanding of parameter space and thus of parametric uncertainty. It will then be possible to generate predictions from models with quantifiable confidence or significance.

Parameter sensitivity and robustness

Sensitivity analysis is a straightforward and informative means to determine which features of a model have the greatest impact on a particular output or model feature. When we evaluated parameter sensitivity in regions of parameter space centered on independent fits (Cukier *et al*, 1973), we observed the rank order of the most sensitive parameters to be relatively constant and thus, relatively unaffected by parametric uncertainty. However, we also found the identities of the most sensitive parameters to be highly dependent on the output being evaluated. Thus, sensitive parameters for ERK versus Akt, or high versus low ligand concentration or HRG versus EGF are frequently distinct. The union of sensitive parameters across all features and conditions encompasses a much larger fraction of the total parameters in the IERMv1.0 model than sensitivity to a single feature. This conclusion contradicts the oft-cited statement that biological pathways are insensitive to variation in parameter values, but is congruent with experience in other fields of dynamical systems theory. Moreover, it emphasizes the fallacy of simply equating 'robustness' of a network with parametric insensitivity. Response duration might be insensitive to changes in parameter *P* but dose–response relationships might be highly sensitive; as we cannot *a priori* claim that only the former is physiologically important, we cannot state that the network is robust to variation in *P*. Where it possible to enumerate all physiologically significant features of the ErbB pathway in development or normalcy versus disease, we speculate that the majority of parameters would prove sensitive under one condition or another. Thus, analysis of robustness is interpretable only with respect to specific physiological behaviors. Behaviors that are common to many cell types and to variation in initial conditions hint at a much more interesting type of robustness in 'design' that we have yet to understand.

Input–output responses of the ErbB network as revealed by modeling and experiment

To ascertain whether IERMv1.0 can predict significant network features not included in the training data, we examined input–output relationships for EGF and pERK or pAkt. Detailed analysis of progesterone-mediated ERK activation in *Xenopus* oocytes (Huang and Ferrell, 1996) has revealed ultrasensitive behavior (Goldbeter and Koshland, 1984) and a high Hill coefficient arising from double phosphorylation of ERK by MEK and from positive feedback (Huang and Ferrell, 1996; Brown *et al*, 1997; Ferrell and Machleder, 1998). We were therefore surprised to find that, over a 10^6 range in EGF concentrations, our model predicted low Hill coefficients (H_{app}=0.32) and experiments confirmed this prediction. As a consequence, ERK was $\sim$20% as highly phosphorylated at 10^{-12} M EGF as at $\sim 10^{-7}$ M (a saturating level). Amplification of signal at low doses of EGF (as implied by $H_{app}<1$) may help to explain why many ErbB-targeted drugs inhibit receptors at significantly lower drug concentrations than they inhibit downstream kinases: a small number of active receptors are sufficient to drive a substantial downstream signal.

The multivariate nature of control over H_{app} in IERMv1.0 makes it challenging to understand mechanistically, but it is

clear that multiple parameter values, particularly those affecting the coupling of Ras, Raf and MEK play a major role in setting dose responsiveness of a receptor-driven MAPK cascade. It should be noted, however, that dose responsiveness is more complex than we have implied hitherto, as it has an important temporal component; Raf activation dynamics, for example, are predicted to change dramatically as EGF concentrations increases (a prediction we have been unable to test due to poor antibody specificity; Supplementary Figure 7D). Further study of these phenomena will require the development of suitable descriptors of time-varying transfer properties in cell signaling cascades. Moreover, it will undoubtedly prove interesting to compare our data-driven approach with recent theoretical work on network biology (Saez-Rodriguez *et al*, 2004; Del Vecchio *et al*, 2008). The clear suggestion is that the coupling of modules under particular conditions (high retroactivity or mismatched impedance at the connections) results in networks, the behavior of which cannot be easily predicted from analysis of the modules in isolation.

Summary and future prospects

Two aspirations motivate the construction and analysis of complex kinetic models: elucidating the roles of individual proteins at the level of specific biochemical reactions and determining how sets of proteins work together to create modules with discrete physiological functions. We have shown progress in both of these areas for a biochemically realistic model of immediate-early ErbB signaling trained against dynamic data. The strong contextual dependency of parameter sensitivity emphasizes that protein activity is meaningful only with respect to a particular physiological function: many parameters that are insensitive under one set of physiologically reasonable conditions are highly sensitive under another. This is also true at the level of network modules. Although the Raf–MEK–Erk signaling cascade has a switch-like input–output behavior in isolation, within the context of the ErbB pathway the apparent Hill coefficient is $\sim$10-fold lower and the response is log-linear. Thus, we cannot generalize from the behavior of a signaling module in isolation to its behavior in a biological network. These are valuable insights with respect to our initial aspirations, but emphasize the importance of developing methods to estimate parameter values and handle parametric uncertainty through optimal experimental design and Bayesian estimation. The challenge is now to implement these concepts in a practical manner with a complex biochemical model.

Materials and methods

Model implementation

The ErbB model was implemented in two software packages: MATLAB and Jacobian. In MATLAB (MathWorks, Natick, MA), the model was constructed using the Simbiology Tool Box, comprising 828 reactions, 499 species and 201 parameters. The model was also reimplemented in Jacobian, a reaction-engineering program from Numerica (Cambridge, MA). The numerical integrator in Jacobian is optimized for sparse systems such as the ErbB model. The model is available in five formats, (i) SBML, (ii) MATLAB Simbiology object, (iii) a set of Jacobian scripts, (iv) a set of text files of reactions, species and parameters or (v) as a set of text files of ODE equations, initial conditions and parameters (http://www.cdpcenter.org/data/Chen_2008, model files). Model dynamics on MATLAB and Jacobian platforms were shown to be consistent.

Sensitivity analysis

Local sensitivity analysis was performed with Jacobian. Sensitivities were fully normalized as in equation (1). Regional sensitivity, defined as an 'average' sensitivity of a parameter around a fit in parameter space, involved repeated rounds of sensitivity calculation and averaging of the value. In each run, all parameters are randomly modulated around the nominal values. We calculated regional sensitivities only in the Jacobian package, because sampling is computationally intensive and Jacobian's numerical solver was able to calculate more samples. Sensitivity averaging and parameter modulation was implemented in Jacobian's scripting environment. For parameter modulation, each nominal parameter was multiplied by a different, random factor of 10^x where x is a random number evenly distributed between -2.5 and 2.5. This gives rise to a model with new parameter values that deviate from nominal ones. The sensitivity of a parameter is calculated for this new model, and entered into the average. This process is repeated a thousand times, and the resulting average sensitivity across randomly parameterized models gives the regional sensitivity.

Simulated annealing

SA was implemented in the scripting environment of Jacobian. The configuration space is given by the vector of parameter values. The 'energy' function is given by the RMSD function that assigns a score to each configuration is the objective function, defined as follows:

$$U = \sum_{i=1}^{N_{\mathrm{exp}t}} \sum_{j=1}^{N_{\mathrm{obs}}} \sum_{k=1}^{N_t} \left(x_{i,j}(t_k) - x^{\mathrm{e}}_{i,j}(t_k)\right)^2$$

where x is the number of activated molecules of some species in the model, x^{e} is the number of activated molecules of the corresponding species in experiment, N_{obs} is the number of measurables or observable species, N_{expt} is the number of experimental conditions and N_t is the number of time points in the experiment. The measured phosphorylation of a species was normalized for each cell line such that the maximum signal for a species among all stimulation conditions was set to 1 (in Figure 2, the signals are further scaled to reach 100). The simulated dynamics for pERK and pAkt were normalized by dividing by the maximal possible activated signal (the initial number of molecules of the unphosphorylated form). For simulated pErbB1, the maximal possible activated signal was assumed to be 70% of the measured number of ErbB1 molecules, as we observed that with higher ligand concentrations, more ErbB1 activation could be observed (data not shown). The normalization conditions for ERK and Akt reflect the assumption that the maximal observed signal is a 'saturating' one, i.e. all species have been phosphorylated. The objective function measures how well a model with a particular configuration of parameters is able to produce dynamics-matching experiment. Of the 201 rate constants and 28 initial protein species numbers, 75 moves are possible (see main text). The move at each iteration of SA in parameter space consists of a multiplicative change in a selected parameter, where the multiplicative factor (or move size) is 10^x where x is a random increment of 0.25. Acceptance of a move is given by the Metropolis criterion, i.e. if the move results in a decrease in the objective function, the move is accepted; if the move results in an increase in the objective function, the probability that a move is accepted is given by $e^{-\Delta E/T}$, where ΔE is the change in objection function value after the move. We set bounds on the moves, so parameters did not deviate more than $10^{2.5}$- or $10^{-2.5}$-fold from the starting point of annealing. These bounds put limits on the size of the explored space. Cooling of the system adhered to an exponential schedule $T{=}T_0\exp(-\mathrm{step}/\mathrm{step}_0)$, where T_0 is the starting temperature, step is the step number in the 4000-step annealing run and step_0 is the time it takes for the temperature to be reduced by a factor $1/e$ (i.e. analogous to the half-life). Through trial

and error, we chose T_0 to be 1, and $step_0$ to be 1700, values that gave the highest number of near optimal fits.

Tissue culture

A431, H1666 and H3255 cell lines were grown in ACL-4 medium, HBL100 and MCF7 were grown in DMEM medium, and T47D and AU565 were grown in RPMI medium. Each medium was supplemented with 10% fetal bovine serum (FBS). Cells were trypsinized and subcultured into 96-well plates (cat. no. 165305; Nalgene Nunc, Rochester, NY, for in-cell western blot assays and cat. no. 3072; Becton Dickinson, Lincoln Park, NJ, for xMAP assays) at varying densities, grown for 24 h and subsequently serum starved for 16 h in the appropriate medium containing 0.1% FBS. At the time of treatment, the confluency of all cell lines was approximately 75%.

Time course measurements by in-cell western blot assay

Time course measurements were performed by in-cell western blot assay. Cells were stimulated by adding ligand for the indicated length of time. At the time of observation, ligand was removed and cells were fixed (4% formaldehyde in 1 × PBS, 20 min at 25°C), permeabilized (0.1% Triton in 1 × PBS, four washes for 5 min each at 25°C with rotation) and blocked (0.5 × Odyssey Blocking Buffer (OBB) (Licor Biosciences, Lincoln, NE), 1 h at 25°C with rotation). Primary antibodies were diluted 1:100 in 0.5 × OBB (50 µl per well) and allowed to incubate overnight at 4°C. Cells were then washed 3 × with 1 × PBS-Tween (0.1%) for 5 min at 25°C with rotation, followed by the addition of secondary antibody (diluted 1:800 in 0.5 × OBB) and TO-PRO-3 (DNA stain, diluted 1:2500 in 0.5 × OBB) and incubation for 1 h at 25°C. Following three additional wash cycles, wells were aspirated dry and scanned using the Licor Odyssey Scanner (Licor Biosciences). Data were normalized on a 'per cell' basis by dividing fluorescence on the 800 channel (pErbB1, pERK or pAkt) by the 700 channel (DNA stain).

IC_{50} and EC_{50} measurements by xMAP assay

EC_{50} and IC_{50} measurements were performed by bead-based immunoassay. For IC_{50} measurements, cells were pretreated for 1 h with the indicated concentrations of Iressa or Lapatinib (LC Laboratories, Woburn, MA). EGF was added at 5 nM for the IC_{50} measurements or at the indicated concentrations for the EC_{50} measurements. After 5-min incubations, the cells were washed and lysed using the Bio-Plex cell lysis kit (Bio-Rad, Hercules, CA). pAkt (S473) and pERK (Y185/Y187) were measured using xMAP bead kits (LHO0101 and LHO0241; Invitrogen, Carlsbad, CA) and pY-ErbB1 was measured using Luminex bead kit (71935-3; EMD Chemicals Inc., San Diego, CA). Assays were performed using the Bio-Plex Phosphoprotein Detection Reagent kit (Bio-Rad) and the Bio-Plex 200 platform (Bio-Rad).

Antibodies

Primary monoclonal antibodies recognizing pErbB1 (Y1068) (cat. no. 2234), pERK 1 (T202/Y204) or pERK2 (T185/Y187) (cat. no. 4377), pAkt (S473) (cat. no. 4058) and MEK (S221) (cat. no. 2338) were obtained from Cell Signaling Technology (Danvers, MA). Primary monoclonal antibody recognizing Shc-Y239/Y240 (cat. no. 44–830) was obtained from Biosource (now Invitrogen). Primary polyclonal antibodies recognizing ErbB1 phosphotyrosine sites Y845 (cat. no. 2237), Y992 (cat. no. 2234), Y1045 (cat. no. 2235), Y1068 (cat. no. 4404) and Y1173 cat. no. 2231 were obtained from Cell Signaling Technology. IRDye 800-conjugated goat-anti-rabbit secondary antibodies were obtained from Rockland Immunochemicals, Gilbertsville, PA, and TO-PRO-3 DNA dye was obtained from Invitrogen.

Supplementary information

Supplementary information is available at the *Molecular Systems Biology* website (www.nature.com/msb).

Acknowledgements

We thank Laura Sontag-Kleiman, Julio Saez-Rodriguez and Carlos Lopez for discussion and editing and Taeshin Park for assistance with Jacobian. This study was financially supported by NIH center grants GM68762 and CA112967.

Conflict of interest

The authors declare that they have no conflict of interest.

References

Aldridge BB, Burke JM, Lauffenburger DA, Sorger PK (2006) Physicochemical modelling of cell signalling pathways. *Nat Cell Biol* **8:** 1195–1203

Amit I, Citri A, Shay T, Lu Y, Katz M, Zhang F, Tarcic G, Siwak D, Lahad J, Jacob-Hirsch J, Amariglio N, Vaisman N, Segal E, Rechavi G, Alon U, Mills GB, Domany E, Yarden Y (2007) A module of negative feedback regulators defines growth factor signaling. *Nat Genet* **39:** 503–512

Bentele M, Lavrik I, Ulrich M, Stösser S, Heermann DW, Kalthoff H, Krammer PH, Eils R (2004) Mathematical modeling reveals threshold mechanism in CD95-induced apoptosis. *J Cell Biol* **166:** 839–851

Berkers JA, van Bergen en Henegouwen PM, Boonstra J (1991) Three classes of epidermal growth factor receptors on HeLa cells. *J Biol Chem* **266:** 922–927

Birtwistle MR, Hatakeyama M, Yumoto N, Ogunnaike BA, Hoek JB, Kholodenko BN (2007) Ligand-dependent responses of the ErbB signaling network: experimental and modeling analyses. *Mol Syst Biol* **3:** 144

Blinov ML, Faeder JR, Goldstein B, Hlavacek WS (2004) BioNetGen: software for rule-based modeling of signal transduction based on the interactions of molecular domains. *Bioinformatics* **20:** 3289–3291

Brown GC, Hoek JB, Kholodenko BN (1997) Why do protein kinase cascades have more than one level? *Trends Biochem Sci* **22:** 288

Brown KS, Sethna JP (2003) Statistical mechanical approaches to models with many poorly known parameters. *Phys Rev* **68** (2 Part 1)**:** 021904

Cao Z, Wu X, Yen L, Sweeney C, Carraway III KL (2007) Neuregulin-induced ErbB3 downregulation is mediated by a protein stability cascade involving the E3 ubiquitin ligase Nrdp1. *Mol Cell Biol* **27:** 2180–2188

Carey KD, Garton AJ, Romero MS, Kahler J, Thomson S, Ross S, Park F, Haley JD, Gibson N, Sliwkowski MX (2006) Kinetic analysis of epidermal growth factor receptor somatic mutant proteins shows increased sensitivity to the epidermal growth factor receptor tyrosine kinase inhibitor, erlotinib. *Cancer Res* **66:** 8163–8171

Carraway III KL, Sliwkowski MX, Akita R, Platko JV, Guy PM, Nuijens A, Diamonti AJ, Vandlen RL, Cantley LC, Cerione RA (1994) The erbB3 gene product is a receptor for heregulin. *J Biol Chem* **269:** 14303–14306

Christie MA, Glimm J, Grove JW, Higdon DM, Sharp DH, Wood-Schultz MM (2005) Error analysis and simulations of complex phenomena. *Los Alamos Sci* **29:** 6–25

Citri A, Yarden Y (2006) EGF–ERBB signalling: towards the systems level. *Nat Rev* **7:** 505–516

Cukier RI, Fortuin CM, Shuler KE, Petschek AG, Schaibly JH (1973) Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. *J Chem Phys* **59:** 3873–3878

Danos V, Feret J, Fontana W, Harmer R, Krivine J (2007) Rule-based modeling of cellular signaling. In *Lecture Notes in Computer Science: Concurrency Theory*, Hutchinson D, Kanade T (eds), Vol. 4703, pp 17–41. Heidelberg: Springer Berlin

Del Vecchio D, Ninfa AJ, Sontag ED (2008) Modular cell biology: retroactivity and insulation. *Mol Syst Biol* **4:** 161

Engelman JA, Zejnullahu K, Mitsudomi T, Song Y, Hyland C, Park JO, Lindeman N, Gale CM, Zhao X, Christensen J, Kosaka T, Holmes AJ, Rogers AM, Cappuzzo F, Mok T, Lee C, Johnson BE, Cantley LC, Jänne PA (2007) MET amplification leads to gefitinib resistance in lung cancer by activating ERBB3 signaling. *Science* **316:** 1039–1043

Felder S, Zhou M, Hu P, Ureña J, Ullrich A, Chaudhuri M, White M, Shoelson SE, Schlessinger J (1993) SH2 domains exhibit high-affinity binding to tyrosine-phosphorylated peptides yet also exhibit rapid dissociation and exchange. *Mol Cell Biol* **13:** 1449–1455

Ferrell Jr JE, Machleder EM (1998) The biochemical basis of an all-or-none cell fate switch in *Xenopus* oocytes. *Science* **280:** 895–898

Gadella Jr TW, Jovin TM (1995) Oligomerization of epidermal growth factor receptors on A431 cells studied by time-resolved fluorescence imaging microscopy. A stereochemical model for tyrosine kinase receptor activation. *J Cell Biol* **129:** 1543–1558

Goldbeter A, Koshland Jr DE (1984) Ultrasensitivity in biochemical systems controlled by covalent modification. Interplay between zero-order and multistep effects. *J Biol Chem* **259:** 14441–14447

Gonzalez OR, Küper C, Jung K, Naval Jr PC, Mendoza E (2007) Parameter estimation using simulated annealing for S-system models of biochemical networks. *Bioinformatics* **23:** 480–486

Graus-Porta D, Beerli RR, Daly JM, Hynes NE (1997) ErbB-2, the preferred heterodimerization partner of all ErbB receptors, is a mediator of lateral signaling. *EMBO J* **16:** 1647–1655

Gutenkunst RN, Waterfall JJ, Casey FP, Brown KS, Myers CR, Sethna JP (2007) Universally sloppy parameter sensitivities in systems biology models. *PLoS Comput Biol* **3:** 1871–1878

Guy PM, Platko JV, Cantley LC, Cerione RA, Carraway III KL (1994) Insect cell-expressed p180erbB3 possesses an impaired tyrosine kinase activity. *Proc Natl Acad Sci USA* **91:** 8132–8136

Hatakeyama M, Kimura S, Naka T, Kawasaki T, Yumoto N, Ichikawa M, Kim JH, Saito K, Saeki M, Shirouzu M, Yokoyama S, Konagaya A (2003) A computational model on the modulation of mitogen-activated protein kinase (MAPK) and Akt pathways in heregulin-induced ErbB signalling. *Biochem J* **373** (Part 2)**:** 451–463

Heinrich R, Neel BG, Rapoport TA (2002) Mathematical models of protein kinase signal transduction. *Mol cell* **9:** 957–970

Hendriks BS, Orr G, Wells A, Wiley HS, Lauffenburger DA (2005) Parsing ERK activation reveals quantitatively equivalent contributions from epidermal growth factor receptor and HER2 in human mammary epithelial cells. *J Biol Chem* **280:** 6157–6169

Hendriks BS, Opresko LK, Wiley HS, Lauffenburger D (2003) Quantitative analysis of HER2-mediated effects on HER2 and epidermal growth factor receptor endocytosis: distribution of homo- and heterodimers depends on relative HER2 levels. *J Biol Chem* **278:** 23343–23351

Hengl S, Kreutz C, Timmer J, Maiwald T (2007) Data-based identifiability analysis of non-linear dynamical models. *Bioinformatics* **23:** 2612–2618

Hlavacek WS, Faeder JR, Blinov ML, Perelson AS, Goldstein B (2003) The complexity of complexes in signal transduction. *Biotechnol Bioeng* **84:** 783–794

Huang CY, Ferrell Jr JE (1996) Ultrasensitivity in the mitogen-activated protein kinase cascade. *Proc Natl Acad Sci USA* **93:** 10078–10083

Ingolia NT (2004) Topology and robustness in the *Drosophila* segment polarity network. *PLoS Biol* **2:** e123

Jones RB, Gordus A, Krall JA, MacBeath G (2006) A quantitative protein interaction network for the ErbB receptors using protein microarrays. *Nature* **439:** 168–174

Kholodenko BN, Demin OV, Moehren G, Hoek JB (1999) Quantification of short term signaling by the epidermal growth factor receptor. *J Biol Chem* **274:** 30169–30181

Kirkpatrick S, Gelatt Jr CD, Vecchi MP (1983) Optimization by simulated annealing. *Science* **220:** 671–680

Klapper LN, Glathe S, Vaisman N, Hynes NE, Andrews GC, Sela M, Yarden Y (1999) The ErbB-2/HER2 oncoprotein of human carcinomas may function solely as a shared coreceptor for multiple stroma-derived growth factors. *Proc Natl Acad Sci USA* **96:** 4995–5000

Kuepfer L, Peter M, Sauer U, Stelling J (2007) Ensemble modeling for analysis of cell signaling dynamics. *Nat Biotechnol* **25:** 8503–8511

Landgraf R, Eisenberg D (2000) Heregulin reverses the oligomerization of HER3. *Biochemistry* **39:** 8503–8511

Mahdavi A, Davey RE, Bhola P, Yin T, Zandstra PW (2007) Sensitivity analysis of intracellular signaling pathway kinetics predicts targets for stem cell fate control. *PLoS Comput Biol* **3:** e130

Margaria G, Riccomagno E, Chappell MJ, Wynn HP (2001) Differential algebra methods for the study of the structural identifiability of rational function state-space models in the biosciences. *Math Biosci* **174:** 1–26

Marquardt D (1963) An algorithm for least-squares estimation of nonlinear parameters. *SIAM J Appl Math* **11:** 431–441

Melke P, Jönsson H, Pardali E, ten Dijke P, Peterson C (2006) A rate equation approach to elucidate the kinetics and robustness of the TGF-beta pathway. *Biophys J* **91:** 4368–4380

Morohashi M, Winn AE, Borisuk MT, Bolouri H, Doyle J, Kitano H (2002) Robustness as a measure of plausibility in models of biochemical networks. *J Theor Biol* **216:** 19–30

Mulloy R, Ferrand A, Kim Y, Sordella R, Bell DW, Haber DA, Anderson KS, Settleman J (2007) Epidermal growth factor receptor mutants from human lung cancers exhibit enhanced catalytic activity and increased sensitivity to gefitinib. *Cancer Res* **67:** 2325–2330

Northrup SH, Erickson HP (1992) Kinetics of protein–protein association explained by Brownian dynamics computer simulation. *Proc Natl Acad Sci USA* **89:** 3338–3342

Ono M, Hirata A, Kometani T, Miyagawa M, Ueda S, Kinoshita H, Fujii T, Kuwano M (2004) Sensitivity to gefitinib (Iressa, ZD1839) in non-small cell lung cancer cell lines correlates with dependence on the epidermal growth factor (EGF) receptor/extracellular signal-regulated kinase 1/2 and EGF receptor/Akt pathway for proliferation. *Mol Cancer Ther* **3:** 465–472

Paez JG, Jänne PA, Lee JC, Tracy S, Greulich H, Gabriel S, Herman P, Kaye FJ, Lindeman N, Boggon TJ, Naoki K, Sasaki H, Fujii Y, Eck MJ, Sellers WR, Johnson BE, Meyerson M (2004) EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. *Science* **304:** 1497–1500

Peles E, Bacus SS, Koski RA, Lu HS, Wen D, Ogden SG, Levy RB, Yarden Y (1992) Isolation of the neu/HER-2 stimulatory ligand: a 44 kd glycoprotein that induces differentiation of mammary tumor cells. *Cell* **69:** 205–216

Resat H, Ewald JA, Dixon DA, Wiley HS (2003) An integrated model of epidermal growth factor receptor trafficking and signal transduction. *Biophys J* **85:** 730–743

Rodriguez-Fernandez M, Mendes P, Banga JR (2006) A hybrid approach for efficient and robust parameter estimation in biochemical pathways. *Biosystems* **83:** 248–265

Rusch V, Baselga J, Cordon-Cardo C, Orazem J, Zaman M, Hoda S, McIntosh J, Kurie J, Dmitrovsky E (1993) Differential expression of the epidermal growth factor receptor and its ligands in primary non-small cell lung cancers and adjacent benign lung. *Cancer Res* **53** (10 Suppl)**:** 2379–2385

Saez-Rodriguez J, Kremling A, Conzelmann H, Bettenbrock K, Gilles ED (2004) Modular analysis of signal transduction networks. *IEEE Contr Sys Mag* **24:** 35–52

Sasagawa S, Ozaki Y, Fujita K, Kuroda S (2005) Prediction and validation of the distinct dynamics of transient and sustained ERK activation. *Nat Cell Biol* **7:** 365–373

Sastry L, Lin W, Wong WT, Di Fiore PP, Scoppa CA, King CR (1995) Quantitative analysis of Grb2-Sos1 interaction: the N-terminal SH3 domain of Grb2 mediates affinity. *Oncogene* **11:** 1107–1112

Schnell S, Turner TE (2004) Reaction kinetics in intracellular environments with macromolecular crowding: simulations and rate laws. *Prog Biophys Mol Biol* **85:** 235–260

Schoeberl B, Eichler-Jonsson C, Gilles ED, Muller G (2002) Computational modeling of the dynamics of the MAP kinase cascade activated by surface and internalized EGF receptors. *Nat Biotechnol* **20:** 370–375

Schulze WX, Deng L, Mann M (2005) Phosphotyrosine interactome of the ErbB-receptor kinase family. *Mol Syst Biol* **1:** 0008

Silverstein AM, Barrow CA, Davis AJ, Mumby MC (2002) Actions of PP2A on the MAP kinase pathway and apoptosis are mediated by distinct regulatory subunits. *Proc Natl Acad Sci USA* **99:** 4221–4226

Singer AB, Taylor JW, Barton PI, Green WH (2006) Global dynamic optimization for parameter estimation in chemical kinetics. *J Phys Chem* **110:** 971–976

Slamon DJ, Clark GM, Wong SG, Levin WJ, Ullrich A, McGuire WL (1987) Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. *Science* **235:** 177–182

Slamon DJ, Leyland-Jones B, Shak S, Fuchs H, Paton V, Bajamonde A, Fleming T, Eiermann W, Wolter J, Pegram M, Baselga J, Norton L (2001) Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. *N Engl J Med* **344:** 783–792

Sordella R, Bell DW, Haber DA, Settleman J (2004) Gefitinib-sensitizing EGFR mutations in lung cancer activate anti-apoptotic pathways. *Science* **305:** 1163–1167

Stein RA, Hustedt EJ, Staros JV, Beth AH (2002) Rotational dynamics of the epidermal growth factor receptor. *Biochemistry* **41:** 1957–1964

Tracy S, Mukohara T, Hansen M, Meyerson M, Johnson BE, Jänne PA (2004) Gefitinib induces apoptosis in the EGFRL858R non-small-cell lung cancer cell line H3255. *Cancer Res* **64:** 7241–7244

Ugi S, Imamura T, Ricketts W, Olefsky JM (2002) Protein phosphatase 2A forms a molecular complex with Shc and regulates Shc tyrosine phosphorylation and downstream mitogenic signaling. *Mol Cell Biol* **22:** 2375–2387

van Riel NA (2006) Dynamic modelling and analysis of biochemical networks: mechanism-based models and model-based experiments. *Brief Bioinformatics* **7:** 364–374

von Dassow G, Meir E, Munro EM, Odell GM (2000) The segment polarity network is a robust developmental module. *Nature* **406:** 188–192

Vyshemirsky V, Girolami MA (2008) Bayesian ranking of biochemical system models. *Bioinformatics* **24:** 833–839

White LJ, Evans ND, Lam TJ, Schukken YH, Medley GF, Godfrey KR, Chappell MJ (2001) The structural identifiability and parameter estimation of a multispecies model for the transmission of mastitis in dairy cows. *Math Biosci* **174:** 77–90

Yarden Y, Sliwkowski MX (2001) Untangling the ErbB signalling network. *Nat Rev* **2:** 127–137

Yun CH, Boggon TJ, Li Y, Woo MS, Greulich H, Meyerson M, Eck MJ (2007) Structures of lung cancer-derived EGFR mutants and inhibitor complexes: mechanism of activation and insights into differential inhibitor sensitivity. *Cancer cell* **11:** 217–227